Development and Validation of Situational Judgement Test to Measure Continuous Learning Competency

Yunita Faela Nisa,1 Jahja Umar2, Puspita Dian Arista3, Helena Widaningrum3, Bahrul Hayat1, Faculty of Psychology UIN Syarif Hidayatullah Jakarta, Indonesia1
Institut Asesmen Indonesia2
PT PLN (Persero) Assessment Center Executing Unit3
yunita.faela@uinjkt.ac.id

Abstract
Measuring continuous learning (CL) competency is beneficial to achieving success for individuals and organizations. For this reason, the availability of a brief but valid instrument is crucial. This study aims to develop a valid instrument to measure the CL using a situational judgment test (SJT). The instrument was developed and validated using respondents of 502 employees at the supervisory level in the electrical industry (with age M=30.13, SD=6.17). For content validity, we first held focus group discussions with supervisors and managers to identify and develop some essential characteristics of the CL. Subject matter experts were involved in writing and reviewing the items. A confirmatory factor analysis was conducted to test the construct validity, and it was found that 12 of 13 items fitted a unidimensional model. All the factor loadings were statistically significant (p<.05). A further test for parallel assumption was done to check the tau-equivalence. This test is vital because raw scores are mostly used in daily practice rather than scaled scores. The results showed that 12 items met the tau equivalent requirement. For further research, scoring for multiple-choice SJTs using the nominal response model (NRM) could be considered since there was a tendency for some of the questions to elicit ambiguous choices.

Keywords: continuous learning, confirmatory factor analysis, IRT, situational judgment test

Abstrak

Kata kunci: continuous learning, confirmatory factor analysis, IRT, situational judgment test
Introduction

Organizations thrive when their members’ learning and development are ongoing. For individuals and organizations, continuous learning competence is a critical aspect of supporting the achievement of goals. This competence increases the knowledge and skills needed to achieve success for individuals and the organization (Maurer & Weiss, 2010; Molloy & Noe, 2009).

Continuous learning, represented by the ability to learn and develop one’s skills, is becoming a core career competency (Hall & Mirvis, 1995). This competency improves performance, and the ability to develop the performance has a competitive value for an organization. Given the importance of continuous learning for both individuals and organizations, the topic is worth considering for developing an instrument to measure it.

Ubiquitous changes in information and aptitude prerequisites on the job make continuous learning a necessity (Ilgen & Pulakos, 1999). In addition, continuous learning is vital for anticipating the aptitudes of inexperienced unemployed people who have been out of work. Looking for extra preparation or learning openings can assist in reemployment ( Wanberg et al., 2002; Leana & Feldman, 1988; Vinokur et al., 2000; Vuori & Silvonen, 2005). Continuous learning is additionally vital in helping more seasoned unemployed people extend their opportunities for employment (Wanberg et al., 2000). Given the recent financial emergency and higher unemployment rate, continuous learning could be a convenient and vital issue for both people in the workforce and unemployed people.

Regarding the dynamics of the organizational environment, individual learners’ attributes can also affect their motivation and desire to learn continuously. Previous research findings have shown that older employees are less interested in job-related learning and development (Maurer, 2001). In addition, experience has led many people to consider continuous learning less important (Schmidt et al., 1986). However, continuous learning is necessary for any employee to enhance one’s job-related experience.

The literature on continuous learning is still fractured (Jain & Martindale, 2012; London & Sessa, 2006): it can be found in resource management, education, psychology, and vocational behavior. There are multiple definitions of continuous learning. For London and Sessa (2006), Continuous learning at the individual level is regularly changing behavior based on a deepening and broadening one’s skills, knowledge, and worldview’ (p. 18). Meanwhile, London and Smither (1999) defined continuous learning as a ‘self-initiated, discretionary, planned, and bold pattern of activities sustained over time to apply or transport knowledge for career development (p. 81). A more generic definition of continuous learning by Tannenbaum (1998) is ‘the process by which individual and organizational learning are fostered on an ongoing basis (p. 438)’. In line with the definitions above, Kluge and Schilling (2003) note that continuous learning is a generally unique concept. They further depict continuous learning by pointing out the following features:

- Continuous learning is a continuous preparation for advancement within the organizational setting. It does not have a clear beginning or conclusion. In any case, it eventually must advantage the individual’s career proficiency and the organization. Continuous learning can be seen as a subset of lifelong learning.
- Continuous learning can be both formal and casual. It can incorporate daily, day-long, and anytime learning. It applies at whatever point a person is deliberately considering, reflecting, or learning.
- Continuous learning happens from the individual to the group to the organizational level, and vice versa for bad habits. At the personal level, it is self-directed; at the group level, it is collaborative; and at the organizational level, it includes providing opportunities and building up structures and forms that support learning.

In relation to the development of the instrument for this research, continuous learning can be characterized by: 1) actively discovering new topics for learning; 2) constantly creating and taking advantage of existing learning opportunities; 3) applying knowledge and newly acquired skills in work and learn through its application (PLN, 2017).
So far, researchers and practitioners have measured continuous learning competence using the assessment centre approach employing various assessment methods (PLN, 2017). They use an aptitude test, combined with the interview method and observational checklist. Mickelson (2001, 2002) measures a continuous learning competency using e-portfolios, and Tanenbaum (1998) uses a Likert scale to measure continuous learning. We have not found the use of situational judgment test (SJT) to measure continuous learning competence. For this reason, the researchers try to create an instrument using SJT to measure continuous learning competence. One of the advantages of SJT is that it provides a work context (scenario) for the items.

**Situational Judgement Tests**

The first widely used SJT was the George Washington Social Intelligence test, in which several solutions to each situation were offered in a multiple-choice format, and only one of which was judged correct (Moss, 1926). During World War II, Army psychologists developed measures to assess soldiers' judgment. These assessments provided scenarios and alternative responses to each scenario. Unlike the Likert-type format, developing a continuous learning competency instrument is designed to measure continuous learning skills and evaluate whether respondents know appropriate learning behaviours across various situations. SJT is a practical work-related assessment that enables us to capture the dynamic of continuous learning people have in their organizations. SJT is a prevalent assessment strategy regularly utilized for employee selection and promotion (Whetzel et.al., 2020). SJT presents test-takers with a series of job-related scenarios revolving around various issues. For each scenario, the test takers are asked to select from a list of options, and the test-takers choices are then assessed. SJT has been used in employment testing for nearly a century (McDaniel et al., 2001).

According to McDaniel and Nguyen (2001), there are two types of response instructions in SJT constituting two distinct categories: knowledge and behavioral propensity. In knowledge response instructions, respondents are asked to select the correct or best possible response or judge the responses' effectiveness. In behavioral tendency response instructions, respondents are asked to choose the response that best represents what they would likely do or rate the likelihood of doing something (McDaniel et al., 2001). For this research, we use the behavioral tendency approach to measure continuous learning competence.

**Methods**

This section describes the test development and test validation. We divide it into a test development subsection, followed by a subsection of test validation.

**Test Development**

In the first stage, to identify work-related experiences in continuous learning competency, we conducted two focus group sessions with groups consisting of six electrical industry supervisors and middle managers as Subject Matter Experts (SMEs). We also identified continuous learning behaviors for the SJT based on a literature review (Tannenbaum, 1998; Eddy et al., 2005). We used this approach to understand better the actions taken by employees constituting the continuous learning competency. The continuous learning behaviors are sometimes misclassified. For example, taking an advanced course to increase a skill may be classified as a continuous learning competency or meeting a learning need. We reviewed various works in the continuous learning literature to fully understand effective behaviors related to continuous learning skills.

We found four points that reflect a continuous learning competence from the FGD results. First, people participate in a pertinent learning involvement such as going to a workshop, getting into the coaching process from a peer, or developing an interest in learning. Second, the learning experiences create new competencies connected to the work. Third, the organization recognizes and rewards people who apply new knowledge and skill. This positive chain of events—learning, application, and recognition—can
enhance peoples' self-efficacy or conviction to uncover hidden aptitudes and perform optimally. They can also develop positive attitudes toward learning experiences and new challenges.

During the focus group discussion, the SMEs created scenarios and behavior that would be regarded as a correct choice and developed two other less acceptable alternatives for the scenario. Then, SMEs engaged in a series of meetings to carefully recheck each situation using the criteria developed by McDaniel & Nguyen (2001), covering length, complexity, and comprehensibility. Complexity reflects the degree to which circumstances are troublesome to test-takers. Comprehensibility deals with how clearly the questions convey the meaning of the situations. Finally, we checked whether the choices of behaviors were correct responses and supported by sufficient observational and theoretical basis.

The SMEs discussed each situation and the appropriate response until they fully agreed. If they did not reach a consensus, the scenario in question was dropped (Motowidlo et al., 1997). This test development process produced 13 scenarios for continuous learning in an organizational context.

**Instrument**

The instrument measuring continuous learning competency consisted of 13 items, developed using four Continuous Learning characteristics. Each item consists of two parts. The first part is the stem section containing the working scenario. The second part contained three behavioral choices for the participant to select. The participants are asked to choose the most appropriate/favorable behavior and the least appropriate/favorable behavior.

The following is an example of the working scenario: ‘I am asked by the department head to hold a meeting once a month to discuss a certain issue. Each member of the department must prepare the material related to the issue. Some of the members commented that the meeting is just a waste of time. My response to the comment is ….‘

1. I’ll leave this comment because they don’t have the goal and mission at work.
2. I’ll explain that the preparation activity before the meeting is vital for self-learning.
3. I’ll explain that the meeting is a forum for them to learn from each other.

**Respondents**

The respondents of this study were 502 supervisors working in the PT. PLN Persero, a State Electrical Company of Indonesia (age M=30.13; SD=6.17). The respondents were spread all over Indonesia. The supervisory work of the respondents varied came from units of accountancy, public relation, safety and security, data services, logistics, billing management, controlling and organization, asset management, legal services, contact center, and partnership and community outreach.

**Test Administration and Scoring**

The researchers used online computer-assisted test administration. The test administration is conducted concurrently in October 2017. This test is one of the PLN’s soft competency assessment kits. The respondents were asked to complete the assessment kit in two hours.

The researchers developed the scoring guide as the following:

<table>
<thead>
<tr>
<th>The most favorable</th>
<th>The least favorable</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>Correct</td>
<td>3</td>
</tr>
<tr>
<td>Correct</td>
<td>Incorrect</td>
<td>2</td>
</tr>
<tr>
<td>Incorrect</td>
<td>Correct</td>
<td>2</td>
</tr>
<tr>
<td>Incorrect</td>
<td>Incorrect</td>
<td>1</td>
</tr>
</tbody>
</table>

Using this scoring guide, the score of the respondents ranged from 13 to 39.
Data Analysis

The researcher employed two stages of data analysis. In the first stage, the researchers used Confirmatory Factor Analysis (CFA) with Weighted Least Square Mean and Variance (WSLMV) estimator to investigate the dimensionality of the construct. The researchers used the chi-square statistic and Root Mean Square Error of Approximation (RMSEA) as a criterion for model fit. In the second stage, the researchers tested the parallelism of the Item Characteristic Curve, indicating that the items have the same discriminating power. The purpose of parallelism testing was conducted so that the raw score could be validly utilized. For this parallelism test, the researchers used the CFA-Item Response Theory (IRT) model with the Maximum Likelihood Robust (MLR) estimator. The data of this research fulfill the minimum requirement of sample size since the total sample of this study is more than 500 people (Akour & Al-Omari, 2013). For both analyses, the researcher used Mplus statistical software. The researchers used Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) as criteria for the model.

Results and Discussion

Table 2 shows the correlation matrix of 13 continuous learning items. As shown in Table 2, we can see that the inter-correlations among items ranged from .10 to .41. However, one item (item 10) had a negative correlation with other items. Looking at the content of the item, we found that item 10 contained an ambiguous statement of behavior. For this reason, the researchers dropped item 10, and we excluded the item in the next analysis.

<table>
<thead>
<tr>
<th>Item</th>
<th>CL1</th>
<th>CL2</th>
<th>CL3</th>
<th>CL4</th>
<th>CL5</th>
<th>CL6</th>
<th>CL7</th>
<th>CL8</th>
<th>CL9</th>
<th>CL10</th>
<th>CL11</th>
<th>CL12</th>
<th>CL13</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL2</td>
<td>.15</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL3</td>
<td>.35</td>
<td>.19</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL4</td>
<td>.20</td>
<td>.11</td>
<td>.25</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL5</td>
<td>.37</td>
<td>.20</td>
<td>.47</td>
<td>.27</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL6</td>
<td>.35</td>
<td>.18</td>
<td>.44</td>
<td>.25</td>
<td>.46</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL7</td>
<td>.29</td>
<td>.15</td>
<td>.37</td>
<td>.21</td>
<td>.39</td>
<td>.36</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL8</td>
<td>.32</td>
<td>.17</td>
<td>.41</td>
<td>.23</td>
<td>.43</td>
<td>.39</td>
<td>.33</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL9</td>
<td>.26</td>
<td>.14</td>
<td>.33</td>
<td>.18</td>
<td>.35</td>
<td>.32</td>
<td>.27</td>
<td>.29</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL10</td>
<td>-.14</td>
<td>-.08</td>
<td>-.18</td>
<td>-.10</td>
<td>-.19</td>
<td>-.18</td>
<td>-.15</td>
<td>-.17</td>
<td>-.13</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL11</td>
<td>.33</td>
<td>.17</td>
<td>.42</td>
<td>.24</td>
<td>.44</td>
<td>.41</td>
<td>.35</td>
<td>.38</td>
<td>.31</td>
<td>-.17</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL12</td>
<td>.19</td>
<td>.10</td>
<td>.24</td>
<td>.13</td>
<td>.25</td>
<td>.23</td>
<td>.19</td>
<td>.22</td>
<td>.17</td>
<td>-.10</td>
<td>.22</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>CL13</td>
<td>.32</td>
<td>.17</td>
<td>.41</td>
<td>.23</td>
<td>.43</td>
<td>.40</td>
<td>.34</td>
<td>.37</td>
<td>.30</td>
<td>-.17</td>
<td>.38</td>
<td>.22</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Sources: Personal data

The results of CFA showed the $\chi^2(52) = 113.794; p$-value $< .001$ and the RMSEA (Root Mean Square Error of Approximation) $= .049$ (90% C.I $= .036 – .061$). Based on the criteria of Chi-square, the researchers found that the data did not fit the model. However, from the RMSEA criteria, the researchers found that data fit the model. Since Chi-square is sensitive to the sample size, the researchers prefer to use the RMSEA criterion to check the model fit. The factor loading of Continuous Learning Competency items is shown in Table 3.
Table 3. Factor Loading for items of Continuous Learning Competency

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loading</th>
<th>S.E.</th>
<th>Est./SE</th>
<th>Two-tailed p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL1</td>
<td>.535</td>
<td>.048</td>
<td>11,052</td>
<td>.000</td>
</tr>
<tr>
<td>CL2</td>
<td>.276</td>
<td>.056</td>
<td>4,957</td>
<td>.000</td>
</tr>
<tr>
<td>CL3</td>
<td>.677</td>
<td>.039</td>
<td>17,224</td>
<td>.000</td>
</tr>
<tr>
<td>CL4</td>
<td>.376</td>
<td>.052</td>
<td>7,266</td>
<td>.000</td>
</tr>
<tr>
<td>CL5</td>
<td>.728</td>
<td>.040</td>
<td>18,292</td>
<td>.000</td>
</tr>
<tr>
<td>CL6</td>
<td>.663</td>
<td>.043</td>
<td>15,498</td>
<td>.000</td>
</tr>
<tr>
<td>CL7</td>
<td>.477</td>
<td>.040</td>
<td>11,842</td>
<td>.000</td>
</tr>
<tr>
<td>CL8</td>
<td>.550</td>
<td>.040</td>
<td>13,662</td>
<td>.000</td>
</tr>
<tr>
<td>CL9</td>
<td>.470</td>
<td>.043</td>
<td>11,010</td>
<td>.000</td>
</tr>
<tr>
<td>CL11</td>
<td>.642</td>
<td>.046</td>
<td>14,090</td>
<td>.000</td>
</tr>
<tr>
<td>CL12</td>
<td>.360</td>
<td>.051</td>
<td>7,070</td>
<td>.000</td>
</tr>
<tr>
<td>CL13</td>
<td>.623</td>
<td>.047</td>
<td>13,216</td>
<td>.000</td>
</tr>
</tbody>
</table>

Sources: Personal data

Table 3 showed that 12 items have a good factor loading ranging from .276 to .728. The researchers concluded that the 12 items are valid for measuring continuous learning competency. Figure 1 shows the factor loading of continuous learning competency items:

Figure 1. CFA Model of Continuous Learning Competency

Figure 1 describes the path diagram of the CFA model of continuous learning competency. All 12 items are unidimensional. We freed up measurement error correlation between items 7 and 9 to generate a fit model, and the results show that the test items are unidimensional. It means that the 12 SJT items do measure continuous learning competency.
The confirmatory factor analysis IRT for the continuous learning competency showed that \( \chi^2(531,179) =14,493.581; \) p-value=1.000. The chi-square test for parallel unidimensional models showed a perfect fit, as shown by the Akaike Information Criterion (AIC) with 11,142.250 and the Bayesian Information Criterion (BIC) with a value 11,247.715. The parameter values are shown in the Table 4.

### Table 4. Factor Loading for items of Continuous Learning Competency

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor loading</th>
<th>S.E.</th>
<th>Est./SE</th>
<th>Two-tailed p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL1</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL2</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL3</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL4</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL5</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL6</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL7</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL8</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL9</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL11</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL12</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
<tr>
<td>CL13</td>
<td>.484</td>
<td>.015</td>
<td>32.326</td>
<td>.000</td>
</tr>
</tbody>
</table>

Sources: Personal data

Table 4 shows that all items have the same factor loading coefficients, which means that each item has the same discriminating power. From the two stages of analysis, the researchers found that the 12 items fit the unidimensional model and fulfilled the tau equivalent requirement. Based on the results, the researchers conclude that a bifactor model analysis is not required. According to Gibbons and Hedeker (1992), a bifactor model is a unidimensional model in which some items empirically measure more than one factor, and there are items that contain a bias (measuring specific factors). For example, a mathematics item written using complicated sentences can measure two dimensions at a time, namely mathematical ability and language ability. As a consequence, without further analysis for a bifactor model, the raw scores can be used for a continuous learning competency test. Following Joreskog (1971), Lord and Novick (1968), and Graham (2006), a test in which all items are valid and have the same factor loading means that each item contributes the same weight to the total score.

Figure 2 showed the results of the CFA-IRT test for the Continuous Learning Competency. It showed that CL items have the same discriminating power.
Figure 2. CFA IRT Model of Continuous Learning Competency

Usually, the model fitness, as shown by the results of this study, is hard to achieve unless the item writers have enough experience and good mastery of the item content (Comrey, 1988; Converse & Presser, 1986). Items are usually vulnerable to bias, coming from the situation/context, culture, gender, diction with multiple interpretations and multiple connotations, or other factors. However, the CL competency items with SJT format have a minimum potential of bias due to the fact that the item scenarios were developed based on real work situations, and the item options reflect choices of plausible behaviors. Another advantage of the SJT format is that it can reduce the effects of social desirability with the real situations faced at work. However, the SJT format can increase the possibility of the multidimensionality of the test data.

The item curves, as shown in Figure 3, indicate that the 12 items are parallel.

Figure 3. Item Characteristic Curve for Continuous Learning Competency Items

Figure 3 shows IRT graphs where each item measuring continuous learning competency has the same discriminating power. This means that all items can distinguish people with low and high continuous...
learning competencies. The probability for people with high continuous learning competency is higher than for those with low continuous learning competency.

**Conclusion**

Based on the two stages of analysis mentioned above, using the CFA procedure, the unidimensionality of the continuous learning competency test is proven. In addition, the CL competency test items have the same discriminating power (tau-equivalent) as shown by the CFA IRT analysis result.

This study is explorative in nature and presents the development procedure and validation in the SJT format to measure continuous learning competency in the electrical industry. The SJT format demonstrated an acceptable psychometric property and is valid for measuring continuous learning competency, especially in the context of the electrical industry.

Other researchers and test developers can take advantage of the result of this study related to continuous learning competency assessment using SJT format and can be applied to other situations measuring various constructs. Researchers can use past empirical studies, theoretical reviews, and meta-analytic findings to identify what specific behaviors and work situations to develop SJT items. Using SJT, the work context written in the items stem will give the respondent a real work context. The work context makes it the respondents easier to make the appropriate choices referring to the scenarios. This can not be done using items with a Likert scale format.

One final concern of this study is related to the external validity of the research. An SJT of continuous learning competency is typical for the Indonesian context; other cultures may have quite different contexts of the type of behavior and work context. This suggests that the factor structure found in this study may not hold across different cultures (Johnston & Hawke, 2002). It is worth noting that using a specific job context in the SJT stem also has limitations in that the item scenarios are not interpreted in the same way by people from different cultures. For example, indicators of the need to learn and competency mastery needed in the job are different across cultures. The results of this research indicated that the raw score could be used for practical purposes. However, we suggest further research using the scoring method with a nominal response model (NRM) as applied by Zu and Kyllonen (2020). They used multiple-choice scoring by selecting one option reflecting the best choice.

**Acknowledgements**

The authors thank PT PLN (Persero) *Pusat Pendidikan dan Pelatihan Unit Pelaksana* Assessment Centre and all the participants for taking part in the study.

**Conflict of Interest**

The authors have no conflict of interest to declare.

**Author Contributions**

YFN, PDA, HW and JU conceived and designed the study. Recruitment, data collection, and data management were done with the assistance of YFN, PDA, and HW; data analysis with the assistance of JU; manuscript draft, including tables and figures, by YFN and BH. All authors reviewed the manuscript, provided comments, and approved the final version.

**Data Availability Statement**

The data that supports the findings of this study are available from the corresponding author upon reasonable request.

http://journal.uinjkt.ac.id/index.php/jp3i

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


Moss, F.A. (1926). Do you know how to get along with people? *Scientific American,* July 1926


