DEVELOPMENT AND PSYCHOMETRIC PROPERTIES OF CAT-BASED COUNSELING ALLIANCE INVENTORY–PEER RATING /OBSERVER

Herdi Herdi*, Sunaryo Kartadinata, dan Agus Taufiq

ABSTRACT
The purpose of this study is to develop and examine psychometric properties of CAT-based Counseling Alliance Inventory–Peer Rating/Observation (CAI-PR/O). To this end, instrument development and standardization procedure was applied. Sixty-one pre-service school counselors from Guidance and Counseling Department in an Indonesian state university were recruited. The data were collected using initial version of CAI-PR/O. Operationally, data were analyzed using Rasch Model version 3.75. The result of the study showed that the developed instrument exhibit satisfying psychometric properties (i.e., item measure, item fit order, DIF, person measure, person fit order, variable maps, rating scale, test reliability, person reliability, and item reliability). However, it was necessary to improve the factor structure. Future studies are recommended to examine the psychometric properties of CAI-PR/O in participants with diverse sociodemographic factors. The result of this study allows researchers to use CAT-based CAI-PR/O as an alternative to measure the counseling alliance accurately. Keywords: computer assisted testing, counseling alliance, counseling alliance inventory–peer rating/observer version, instrument’s development and standardization, Rasch model

INTRODUCTION
Counseling alliance emerges as an essential common factor, as well as a key variable of each counseling approach (Nelson-Jones, 2013; Moss & Glowiak, 2013). It serves as pivotal, fundamental determiner (Norcross J. C., 2010; Norcross & Lambert, 2011a; 2011b; Norcross & Wampold, 2011; Imel & Wampold, 2008) that drives and influences the success of counseling process (Lambert & Vermeersch, 2008; Mozdzierz, Peluso, & Lisiecki, 2008; Lambert, Maximizing psychotherapy outcome beyond evidence-based medicine, 2017). Horvath (1979; Dryden, 2008; Fluckiger, DelRe, Wampold, Symonds, & Horvath, 2012; Fluckiger, Del Re, Wampold, & Horvath, 2018) defines it as a collaborative
relationship between a counselor and the client, which is indicated by emotional attachment and agreement upon the purpose and tasks of a counseling process. Counseling alliance is a conscious, purposeful aspect that involves counselor-client collaboration that is built on active commitment toward specific responsibilities and active, enthusiastic involvement during the counseling process.

In the last two decades, myriad of literature prove that counseling alliance is a pivotal component and a robust key predictor of counseling outcomes (Duff & Bedi, 2010; Horvath, The alliance, 2001). A counselor’s ability in determining, developing, and maintaining alliance with his/her client is proven to be a significant factor of the client’s positive changes (Orlinsky, Ronnestad, & Willuzki, 2004; Crits-Christoph, Gibbons, & Mukherjee, Process-outcome research, 2013; Norcross J. C., 2010). A meta-analytical study on 295 works that involve 30,000 clients shows a positive, significant relationship (categorized as moderate) (Fluckiger, Del Re, Wampold, & Horvath, 2018) with effect size ranging from .21 to .29 (Hardy, Cahill, & Barkham, 2007; Horvath & Bedi, The alliance, 2002) between counseling alliance and higher counseling technique outcomes (Hardy, Cahill, & Barkham, 2007).

Counseling alliance, in few first sessions, serves as the “window of opportunity” in a counseling process and outcome (Bachelor & Horvath, 1999). When a client deems the counseling alliance is well-established during the early sessions (usually in third session), the counseling process tends to result in positive outcome (Fluckiger, Del Re, Wampold, Symonds, & Horvath, 2012; Fluckiger, Del Re, Wampold, & Horvath, 2018; Wampold, 2010; Crits-Christoph P. , Gibbons, Hamilton, Ring-Kurtz, & Gallop, 2011). It is found that positive counseling alliance established during the third session may improve the counseling outcomes by 67% (Stargell, 2017). In contrast, the counseling process may be ended prematurely when the client views that the alliance is poor (Castonguay, Constantino, & Holtforth, 2006). Consequently, counselors are obliged to develop, establish, and maintain a positive alliance since the first session of the counseling. In addition, they are required to be able to manage and repair alliance rupture since it positively affects (ES = .24) the counseling outcome (Safran, Muran, & Eubanks-Carter, 2011).

One of the early efforts to understand, examine, and develop the counseling alliance could be done by assessing the alliance quality of prospective counselors. However, to date, there is no intensive study conducted to develop a standardized instrument for measuring counseling alliance, let alone those to depict the profile of counseling alliance in university level, especially in Jakarta State University.

Horvath and Greenberg (1986; 1989) have developed an instrument to measure the counseling alliance for observers, namely Working Alliance Inventory-Observer which has been widely used and modified into a short form by several scholars such as Falkenstrom, Hatcher, Skjulsvik, Larsson, & Holmqvist, 2015; Hatcher & Gillaspy, 2006; Kokotovic & Tracey, 1989; Murder, Wilmers, Leonhart, Linster, & Barth, 2010). To conclude, WAI-O represent a valid, reliable, and widely-used alliance instrument that has been used in various field, population, and counseling approaches (Paap & Dijkstra, 2017; Sturgiss, et al., 2018; Santirso, Martín-Fernandez, Lila, Gracia, & Terreros, 2018; Vohringer, et al., 2013). Moreover, this instrument has been adopted and translated to 18 languages around to world for a range of purposes (http://wai.profhorvath.com). Since the form of this instrument is still paper-and-pencil test, its administration is carried out manually. Studies that adopt this instrument in online counseling settings are still limited (Penedo, et al., 2019; Miragall, Banos, Cebolla, & Botella, 2015). In Indonesian context, intensive study for developing a standardized instrument to measure and analyze the counseling alliance of prospective school counselor from peer/observer rating has not been reported.

Grounded from the problems described earlier, it is necessary to develop a
standardized Computer-Assisted Testing (CAT)-based counseling alliance instrument for prospective counselors in peer and observer versions. Accordingly, in this study, we developed such instrument called Counseling Alliance Inventory-Peer Rating/Observer (CAI-PR/O). An assessment expert highlights the importance of developing, using, and optimizing CAT system in psychological assessment. CAT system may significantly influence the assessment process, allow easier dissemination of information and wider access to result-based information, offer more accurate result, which eventually results in faster, more efficient, and more accurate data processing, interpretation, and presentation (Drummond & Jones, 2010; Hays, 2013; Whiston, 2009).

The present study focuses on the development and standardization of CAI-PR/O. The novelty of this study lies in its CAT-based product, i.e., CAI-PR/O that is developed based on the most recent theory of Integrative framework, which suggest four dimensions of counseling alliance namely, bond, tasks, goals, and views (Dryden, 2006; 2008).

**RESEARCH METHODOLOGY**

To achieve the purpose of the study, instrument development and standardization procedure was applied. This method was applied to develop and to test the psychometric property of CAI-PR/O. To validate the instrument, an expert and a practitioner in guidance and counseling field were recruited. To test the readability of the instrument, eight guidance and counseling department students were involved. The participants of the field test were 61 guidance and counseling department students of UNJ. They represented students who had not, were joining, and had joined counseling theory and practice courses. CAI-PR/O, as the result of initial development process, was employed to collect data related to peer/observer version of counseling alliance. Using 5-point likert scale, (1 = extremely unlikely, 5 = extremely likely) consisted of 24 items that measured four dimensions of counseling alliance, i.e., bond, tasks, goals, and views. The item feasibility test was done using item measure, item fit order, and differential item functioning analyses. The instrument feasibility test was done using unidimensionality, rating scale, item-person maps, exploratory and confirmatory factor analyses. The reability tests (item reliability, person reliability, and test reability) were done using Cronbach’s Alpha. The collected data were processed using software Rasch Model v. 3.73.

**RESULT AND DISCUSSION**

**Development of Alliance Instrument**

**Measurement**

**Concept Operationalization.** Conceptually, counseling alliance is defined as “the quality and strength of collaborative relationship between a counselor and his/her client, which is indicated by emotional attachment and agreement with tasks and purposes of the counseling.” Counseling alliance covers following four dimensions: Bond, this dimension refers to emotional attachment between counselor and the client that emphasize mutual trust, acceptance, and care in the counseling activity. Tasks, it deals with the agreement between counselor and his/her client regarding various cognitive and behavioral tasks that form the substance of counseling process. Goals, this dimension refers to the agreement between the counselor and his/her clients related to the outcomes of the counseling process (Bordin, 1979; Horvath, 1994; 2001; Horvath & Bedi, 2002; Horvath & Greenberg, 1989). (Dryden, 2006; 2008) comes with the fourth dimension, i.e., View, that refers to the level of similarity of view on various issues.

Operationally, counseling alliance is described as the total score of emotional attachment, agreement in tasks and goals of counseling, and similarity of views on counseling issues between a counselor and his/her clients based on peer/observer rating using CAI-PR/O.

**Dimension and Indicator.** Dimensions and Indicators of counseling alliance are displayed in Table 1 below.

Table 1
Item construction. Each indicator was represented by six items, thus, there were 24 items in total. Five responses range were provided, from 1 = extremely unlikely 2 = unlikely; 3 = fairly likely, 4 = likely, 5 = Very likely.

Analyzing the content and Conducting Pilot testing, revision, and Administering items. Content analysis/ validation of CAI-PR/O was conducted in September 2019 by an expert and a practitioner in guidance and counseling field. They provided a general judgment on the instrument, as displayed in Table 2.

The expert and practitioner judged the items based on the available rating options, namely 1 (less readable/poor/ less suitable), 2 (fairly readable/ fairly good/ fairly suitable, or 3 (readable/good/ suitable).

Table 2
Result of Expert and Practitioner Judgment

<table>
<thead>
<tr>
<th>Aspect</th>
<th>P₁</th>
<th>P₂</th>
<th>Score</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readability</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>Good/understood</td>
</tr>
<tr>
<td>Item suitability with indicators</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>Good/understood</td>
</tr>
<tr>
<td>Item suitability with response patterns</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>Good/understood</td>
</tr>
<tr>
<td>Total average</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>Good/understood</td>
</tr>
</tbody>
</table>

According to the expert and the practitioner, the items of CAI-PR/O had already been readable, suitable with the indicator and the response patterns. They suggested to change the term “negotiate” in item S₂₁ with more common terms. In general, the expert and practitioner concluded that, after revision, CAI-PR/O is feasible.

Pilot testing was conducted in 17 September 2019 by involving eight preservice school counselors from Guidance and counseling department of UNJ. The result showed that the readability of CAI-PR/O items were categorized as good, with the average score of 2.61 out of 3. They also suggested to revise item S₂₁, particularly the word “menegosiasikan” since it is not too common. Overall, the pilot testing concluded that CAI-PR/O is feasible to use.

Pilot testing CAI-PR/O was pilot tested in 18 September to 18 October 2019. The participants of the pilot testing were 61 pre-service school counselors from Guidance and Counseling department of UNJ. The number of the participants of this pilot study is considered enough to obtain a stable result within the scale of ± 1 logit (Linacre, 2019; Sumintono & Widhiarso, 2014).

Following the pilot testing, the psychometric properties of the instrument were tested. The psychometric properties test included item accuracy (i.e., item measure, item fit order, and differential item functioning (DIF)), respondents’ abilities (i.e., person measure, person fit order, and item-person maps), quality analysis (unidimensionality and rating scale), and reliability (test reliability, person reliability, and item reliability). The test was done using Item Response Theory (IRT) approach with Rasch Model using Winsteps 3.73.

The feasibility of CAI-PR/O items were analyzed based on the criteria of item emasure, item fit order, and DIF using Rasch Model. Item measure was used to measure the difficulty of items. The test result showed that item S₁₁ (+1.52 logit) was the most difficult item to be agreed by the participants, while item S₄ (-1.28 logit) was the easiest item to be agreed on.

Item fit order was used to measure identity fit and misfit items. The result indicated that 20 items were fit. The decision
was made based on the criteria that an item is considered fit if the Infit and Outfit Mnsq values are within the range of acceptable value (.5 MNSQ < 1.5), Outfit Zstd (-2 < Zstd > +2), Pt-MCorr (.4 < pt-MCorr .85) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; Sumintono & Widhiarso, 2015; Boone, Staver, & Yale, 2014). Misfit items were S6, S11, S23, dan S24.

DIF functions to detect item bias in certain participant categories. An item is detected as bias based on the probability value of less than 5% (.05) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; Sumintono & Widhiarso, 2015). The test result showed that all items were not biased to certain gender.

Based on these three criteria, 19 items were considered fit. To balance the number of each items for each indicator, sixteen items were selected. They were: S1, S2, S3, S5, S7, S8, S9, S16, S17, S18, S19, S20, S21, dan S22. The distribution of items in final CAI-PR/O is displayed in Table 4.4. Items S10, S14, and S15 were dropped since they had been represented by other items.

The respondents’ ability was analyzed based on criteria of person measure, person fit order, and item-person maps using Rasch Model. The result of person measure test exhibited that respondent no. 20 (+4.65 logit) tend to have high counseling alliance since he gives more answers of “suitable” and very suitable, while respondent 18 (-.08 logit) tend to have low counseling alliance since he mostly answer “not suitable” for each item in CAI-PR/O. Person fit order was used to measure fit and misfit respondents. The criteria is the same as that of item fit order. Based on the criteria, 13 respondents were misfit.

Unidimensionality was used to measure the quality of the instrument, it was done using Rasch Model with principle component analysis from the residue, i.e., measuring the uniformity of the instrument in measuring what suppose to be measured (Linacre, 2019; Sumintono & Widhiarso, 2014; Sumintono & Widhiarso, 2015). The measurement result showed raw variance 37.9% and unaccountable variance of 17.7%.

This value indicates that the minimum unidimensionality threshold of 20% was met, yet the unaccountable variance threshold of ≥ 15% was not met (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

Rating scale criteria was a test to verify the rating options used in the instrument, whether or not it confuses the participants (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). Rasch model analysis found that the average observation score range from logit -3.77 for response 1 (very not likely), logit -1.67 for response 2 (unlikely), logit -1.72 for response 3 (fairly likely), logit +1.27 for response 4 (likely) and logit +4.17 for response 5 (very likely). Similar result was also indicated by Andrich threshold that examine the poltomy used. The value showed a movement from NONE to negative (-2.46, -72) and move to positive direction (.25 .3.03) sequentially. It means that the rating scale used in the instrument has been valid for participant.

Reliability of the instrument is displayed in table 4. The average value of higher than logit .0 indicate that participants tend to choose “suitable” for each item. The Cronbach’s alpha was .84, which is categorized as very good. Meanwhile, the person reliability of .85 and item reliability of .95 indicate that participants’ consistency of answer was good and the items are excellent (Sumintono & Widhiarso, 2014; 2015). This is also supported by Infit and Outfit Mnsq for person and item, where the score was close to the ideal score of 1.00 (i.e., 1.04 and 1.00 and 1.01 and 1.00, respectively). The Infit and Outfit Zstd also shows the average score of person and item of -1 and -2 and 1 and 0, which was considered good since it is close to 0.

<table>
<thead>
<tr>
<th>Table 4 CAI-PR/O Statistical Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Output</strong></td>
</tr>
<tr>
<td>Item</td>
</tr>
<tr>
<td>Separation index</td>
</tr>
<tr>
<td>Separation strata (H)</td>
</tr>
<tr>
<td>Higher logit value</td>
</tr>
<tr>
<td>Lower logit value</td>
</tr>
</tbody>
</table>
Separation value is used to categorize person and item into groups. Accordingly, separation value indicate the quality of an instrument since it identifies the respondents and the item group. The equation used was:

\[ H = \frac{\text{Separation strata}}{2} = 3.48 \text{ (rounded as 3)} \]

This result means there are three groups of respondents.

Finalization and designing guideline was done by considering the result of feasibility test, thus, 16 items were retained for CAI-PR/O. CAI-PR/O is expected to be the assessment tool to measure counseling Alliancing of prospective school counselors from peer/observation rating that is more efficient, practical, and valid. CAI-PR/O measures four dimensions of counseling alliance, namely bond, tasks, goals, and views, which each dimension was measured using three items. The outcome of the instrument is in the form of summed rating with 5-point likert Scale. The blueprint of the CAI-PR/O is displayed in table 5, the complete version of the instrument is in the appendix.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dimension</th>
<th>Item number</th>
<th>Σ item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counseling Alliance</td>
<td>Bond</td>
<td>1, 2, 3</td>
<td>5, 4</td>
</tr>
<tr>
<td></td>
<td>Tasks</td>
<td>7, 8, 9</td>
<td>12, 4</td>
</tr>
<tr>
<td></td>
<td>Goals</td>
<td>13, 16, 17, 18</td>
<td>4, 21</td>
</tr>
<tr>
<td></td>
<td>Views</td>
<td>19, 20, 21, 22</td>
<td>16</td>
</tr>
</tbody>
</table>

The profile of counseling alliance is analyzed using average and percentage. The scoring ranges from 16-18 for total CAI-PR/O and 4-20 for each dimension (i.e., bond, tasks, goals, and views). Classification of counseling alliance, both in a whole and per dimensio, is made based on this criteria:

\[ \bar{\text{X}} < 3 = \text{less satisfying, } \geq 3 \leq 4 = \text{fair, dan } \bar{\text{X}} > 4 = \text{satisfying. Higher score reflects better level of counseling alliance.} \]

<table>
<thead>
<tr>
<th>Table 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAI-PR/O Scoring Guideline</td>
</tr>
<tr>
<td>Very suitable</td>
</tr>
<tr>
<td>Positive</td>
</tr>
<tr>
<td>Negative</td>
</tr>
</tbody>
</table>

**CONCLUSION**

The present study has developed CAI-PR/O with satisfactory psychometric properties (based on the item feasibility, person feasibility, and instrument feasibility). The final CAI-PR/O comprised sixteen items that measure four dimensions of counseling alliance, namely bond, tasks, goals, and view. Each dimension was measured using four items. The result of this study allows researchers to use CAT-based CAI-PR/O as an alternatives to measure the counseling alliance accurately. Future studies are recommended to test the instrument using confirmatory factor analysis on wider participants, either pre-service school counselor, counselor, or field supervisor by considering the sociodemographic proportion.

**ACKNOWLEDGMENT**

We would like to thank the Dean of Faculty of Education of UNJ for funding this research. We would also like to thank the participants, editor, reviewer, and other parties for their contribution and help during the process of this research and publication.

**REFERENCES**


Bond, T. G., & Fox, C. M. (2015). *Applying the rasch model: fundamental*
Horvath, A. O., & Greenberg, L. S. (1986). The development of the working alliance inventory. In L. S. Greenberg,


Dryden, & A. Reeves (Eds.), Key issues for counseling in action (pp. 1-17). London: Sage Publications.


ilmu-ilmu sosial (2nd ed.). Cimahi: Trim Komunikata.