The Development and Validating of the New Wisdom Scale for Multicultural Counselor

Herdi Herdi, Sunaryo Kartadinata, Agus Taufiq, aFaculty of Educational Sciences, Universitas Negeri Jakarta, Jakarta, 13220, DKI Jakarta, Indonesia, b,cFaculty of Educational Sciences, Universitas Pendidikan Indonesia, Bandung, 40154, West Java, Indonesia, Email: aHerdi@unj.ac.id, herdiunj5@student.upi.edu

Wisdom is a fundamental personality characteristic and is indicative of a multicultural counselor’s competence. Therefore, an instrument that accurately measures the wisdom of a multicultural counselor is necessary. This purpose of study was to develop a valid and reliable measurement tool to determine the wisdom of multicultural counselor candidates. This study employed design-based research and instrument development procedure. The participants consisted of 517 counselor candidates who were students in the sixth semester of eleven Guidance and Counseling Departments in Indonesian higher education institutions. The data were collected by using an initial questionnaire called New Wisdom Scale for Multicultural Counselors in the form of Likert five-level scale. Operationally, the data analysis was performed by using Rasch Model version 3.75. The results demonstrated that the instrument had satisfying psychometric properties, (i.e. item measure, item fit order, DIF, person measure, person fit order, item-person maps, rating scale, test reliability, person reliability and item reliability). However, the factor structure is still needed to be improved. Further studies are needed to examine the psychometric properties of instruments in various participants.

Key words: Assessment, Counselor Education and Supervision, Multicultural Counselor, Psychometric Properties, Wisdom, Wisdom scale, Rasch Model.
Introduction

Excellent and professional performance of multicultural counselors demands more than intellectual intelligence and multicultural counseling competences but also the need to have wisdom (Hanna, Bemak, & Chung, 1999; Levitt & Piazza-Bonin, 2014; Phan, Torres-Rivera, Volker, & Maddux, 2009). Wisdom appears as a fundamental quality of personality and is the peak competency of effective multicultural counselors (Hanna, Bemak, & Chung, 1999; Levitt & Piazza-Bonin, 2014; Phan, Torres-Rivera, Volker, & Maddux, 2009) and a form of the main issue of multicultural counseling today - ethics virtue (Durodoye, 2013). Wisdom is an application of tacit knowledge directed by values to achieve general benefit (Karelitz, Jarvin, & Sternberg, 2010; Sternberg, 2001; Sternberg & Hagen, 2019). Hanna, Bemak, and Chung (1999) interpret wisdom as a set of cognitive and affective traits and awareness that are needed by counselors to navigate behavior, to choose, to measure, to make complex decisions that are accurate, to overcome conflict and to adapt effectively in complex and uncertain multicultural life.

Previous studies report that counselors’ wisdom provides a significant positive effect of 14% on multicultural counseling competencies (Phan, Torres-Rivera, Volker, & Maddux, 2009), 11.4% -35% of counseling alliances (Osterlund, 2014), leadership (Yang, 2011), weighing and making decisions on various immoral vs. moral dilemmas and intrapersonal vs. interpersonal conflict (Thomas, et al., 2019), contribute positively (9.8%) to peace of mind and (27.4%) to peaceful behavior (Herdi, Kartadinata, & Taufiq, 2017). Other studies found that wisdom is positively correlated with quality of life for individuals and multicultural communities, such as mental health (Roharikova, Spajdel, Cvikova, & Jagla, 2013; Webster, Westerhof, & Bohlmeijer, 2012; Webster, Westerhof, & Bohlmeijer, 2014), happiness (Bergsma & Ardelt, 2012; Etezadi & Pushkar, 2013; Weststrate & Gluck, 2017; Zacher & Staudinger, 2018), life satisfaction (Ardelt, Landes, Gerlach, & Fox, 2013; Shi, Ardelt, & Orwoll, 2016), psychological well-being (Ardelt, 2016; Ardelt & Edwards, 2015; Ardelt & Jeste, 2016; Zacher & Staudinger, 2018), coping strategies, self-control, positive involvement of life for a happy life (Etezadi & Pushkar, 2013), health, hope, self-esteem, positive feelings, humility, social relations and life satisfaction (Krause, 2016; Krause & Hayward, 2014; 2015). In essence, wisdom can guide individuals to transform life and to live right and have better lives (Yang, 2013; Sambo, Aghojare & Ahutu 2016).

The importance of the wisdom of a multicultural counselor is needed to address the challenges of the 21st century and the future of society as they become plural and multicultural. This condition always confronts multicultural counselors in dilemmas and conflicts of intrapersonal, interpersonal, and personal interests, as well as between certainty and uncertainty. Wisdom is also important because it can reduce social scales and negative attitudes toward one-self and others (Webster D. J., 2010). It overcomes the sources of
multicultural counseling barriers such as social class differences and cultural values, stereotypes, cultural biases, encapsulation, resistance and lack of cultural relativism (Baruth & Manning, 2016). Wisdom enable counselors to be ethically responsible to confront various prejudices and discriminations of individuals, groups and communities as well as racism, sexism, classism, homo-prejudice, ableism, ageism and spiritual and religious bias (Remley & Herlihy, 2016). Conversely, the absence of wisdom makes the “cultural encounter” between counselors and counselees to stereotypes, cultural bias, ethnocentrism, encapsulated, and conflicts of interest. Multicultural counselors can also be trapped into the behavior of “foolishness” (Sternberg, 2005; 2019; Aczel, 2019), which can endanger, inhibit, and even lead to failure of the process of alliances and outcomes of multicultural counseling.

Based on the above, the existence of a standardized measuring instrument to measure the wisdom of a multicultural counselor is necessary. The instrument can be a valuable tool in the operationalizing meaning, mapping profiles, designing programs to develop wisdom, to identify and to test the constructs of ontogenesis and the consequences of wisdom, as well as individual differences in wise multicultural counselors (Hanna, Bemak, & Chung, 1999; Karelitz, Jarvin, & Sternberg, 2010). However, until now, there has been no intensive study to develop standardized instruments to measure the wisdom of multicultural counselors. Measuring instruments have tended to be aimed at measuring personal wisdom such as Three-Dimensional Wisdom Scale (Ardelt, 2003), Self-Assessed Wisdom Scale (Webster, 2003; 2007), Wisdom Development Scale (Brown & Greene, 2006; Greene & Brown, 2009), Adult Self-Transcendence Inventory (Levenson, Jennings, Aldwin, & Shiraishi, 2005) and Wise Thinking and Acting Questionnaire (WITHAQ; Moraitou & Efklides, 2012). The purpose of this study is to develop and to test psychometric properties of measuring the wisdom of multicultural counselor’s candidates. This developed measuring instrument uses different constructs from existing instruments, analyzed by using the Item Response Theory (IRT). It is innovative since no previous studies that have ever developed it.

**Method**

**Research Design**

This study uses the procedure of development and standardization of instruments (Heppner, Wampold, & Kivlighan, 2008). This method was chosen because it was intended to develop and to test the psychometric properties of the New Wisdom Scale for Multicultural Counselors (MC-WiSe).

**Research Participants**

The research participants consisted of 517 prospective counselors who enter the sixth semester from eleven higher institutions in Indonesia. Distribution of participants included 135 (26.1%) men and 382 (73.9%) women, 159 (30.8) late adolescents and 358 (69.2) young adults, and 427 (82.6%) Muslims, 34 (6.6%) Catholics, 33 (6.4%) Christians, and 23 (4.4%)
Hinduism. The research participants were selected by using the two stage random sampling technique.

**Measures**

The New Wisdom Scale for Multicultural Counselors (MC-WiSe) is used to measure the wisdom of multicultural counselor’s candidates. The MC-WiSe is developed based on the construct of the wisdom of multicultural counselors (Hanna & Ottens, 1995; Hanna, Bemak, & Chung, 1999). The initial version of MC-WiSe was 68 items to measure 2 dimensions and 11 indicators of the wisdom of multicultural counselors. Cognitive dimensions measured meta-cognition, sharpness of view, dialectical reasoning, tolerance for ambiguity, discovery and alleviation of problems and coping skills. Affective and awareness dimensions measured empathy, caring, recognizing feelings, de-automation and ingenuity. The MC-WiSe takes the form of a five-level Likert scale, starting from 1 = very irrelevant, to 5 = very relevant.

**Research Procedures**

The research procedure to develop the standard instrument of multicultural counselor wisdom is as follows. The first to the third stage is the review of literature, operationalizing the measuring construct, identifying indicators, compiling items, and setting the response format. The MC-WiSe is developed based on the construct of the wisdom of multicultural counselors (Hanna, Bemak, & Chung, 1999; Hanna & Ottens, 1995). Wisdom is defined as “a set of cognitive and affective traits and awareness needed in interpersonal interaction,” living and multicultural counseling effectively. In this initial stage, 68 items were prepared. The item distribution in the initial and the final development stages is presented in Table 1. The response pattern uses a five-level Likert Scale, starting from 1 = very irrelevant to 5 = very relevant.

The fourth stage is the analysis of contents and pilot testing, revision and administering items. At this stage, three experts were engaged from the fields of psychometrics and education evaluation, educational psychology and guidance and counseling. Experts provided an assessment and input on the suitability of items and indicators, the suitability of items with response patterns, and grammar. The fifth stage was a trial. Data collection was carried out traditionally by using a paper and pencil questionnaire by visiting prospective counselors directly in their respective institutions according to the schedule agreed upon by the researcher, the chairpersons of the Study Program, and the participants. The researcher asked for willingness and submitted a guarantee of confidentiality of data to the participants before completing the MC-WiSe. The participants were asked to fill the MC-WiSe in accordance with the instructions. After the data was collected and verified, 517 complete data was obtained for further analysis. The final stage tested psychometric properties and then finalized the items. Various tests to obtain evidence of the validity and reliability of MC-WiSe instruments were carried out at this stage.
Data Analysis Technique

The data analysis techniques included various tests of item accuracy (item measure, item fit order, and differential item functioning [DIF]), person's ability (person measure, person fit order, person maps), analysis of instrument quality (unidimensionality and rating scale), and reliability (test reliability, person reliability, item reliability). The test used the Item Response Theory (IRT) approach, namely the Rasch Model with Winsteps 3.73.

Results and Discussion

The Item Fit

The first question about the description of the item fit of the MC-WiSe item was analyzed from the criteria of item measure, item fit order and DIF using the Rasch Model. The Item Measure was used to measure the difficulty level of items. The test results highlighted the X35 item with +1.10 logit as the item that was most difficult to be approved by the participant, while the item X38 with the value -1.09 logit was the easiest item to approve. These results are visualized in Graphic: Item-Person Maps.

Item fit order was used to determine if the item was a fit or misfit. The results of the study found 31 of 68 items were a fit. This decision was based on the criterion that the item is said to be fit if the Infit and Outfit values of Mnsq are in the range of acceptable values (.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; 2015; Boone, Staver, & Yale, 2014). In this study, the Zstd Outfit criteria were not used because of the large sample size (>500 people) (Sumintono & Widhiarso, 2014; Sumintono & Widhiarso, 2015). The fit items are X1, X3, X5, X6, X7, X8, X11, X13, X14, X17, X25, X26, X27, X30, X37, X38, X41, X47, X48, X49, X50, X51, X52, X55, X57, X59, X61, X62, X64, X67, and X68.

DIF is used to detect item bias in certain participant categories. The detected items are known to be based on a probability value of less than 5% (.05) (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). The test results highlighted 10 items that were biased towards particular sex, namely X19 (.008), X22 (.01), X28 (.003), X32 (.02), X33 (.02), X35 (.000), X37 (.01), X53 (.01), X54 (.008), and X61 (.006). Seven items were biased towards a certain age, namely X2 (.007), X6 (.009), X28 (.001), X39 (.02), X47 (.04), X54 (.02), and X56 (.001).

Based on the three criteria, 28 of 31 items were fit. For the sake of balancing items on each MC-WiSe indicator, 28 items were selected. The final MC-WiSe items were X1, X6, X7, X11, X13, X17, X18, X19, X22, X25, X27, X30, X32, X37, X38, X41, X49, X50, X52, X55, X57, and X64. The item distribution on MC-WiSe indicators and dimensions is presented in Table 1. Items X3,
X5, X8, X14, X26, X48, X51, X59, X62, and X68 were not selected even though they were fit as other items already represented them. Items X18, X19, X22, X32, and X41 remain selected for the balance of items in and the representation of indicators. The same considerations also applied to items X38 even though they were considered most easily approved by respondents because they met the three fit criteria. This is consistent with the opinion (Sumintono & Widhiarso, 2014; 2015) that items that only meet one fit criterion can still be used if it is needed. Other considerations are based on the results of testing the item-total correlation from the Classical Test Theory (CTT) approach by using IBM SPSS 25.0 for Windows for the items X18, X19, X22, X32, X38, and X41 as they have a coefficient index $r_{it} \geq .300$. Psychometric experts say that the rhythm coefficient index is acceptable (Azwar, 2011; Drummond & Jones, 2010).

**Person’s Ability**

The second question about the description of respondents’ ability level evidence was analyzed by the criteria of person measure, person fit order and item person maps by using the Rasch Model. The Person Measure test results show the respondent number 498 (8LDPI) with +1.82 logit tended to have high wisdom because the answers were mostly relevant and very relevant. On the other hand, respondents number 40 (0PDID) with a value of -.40 logit tended to have low multicultural wisdom since the answers were mostly irrelevant. The Person Fit Order is used to test respondents who are fit and misfit. The criteria used are the same as analysis of order fit items. Based on these criteria, 87 of 517 respondents were misfit. Interactions between items and person are visualized in Figure 1. Item-Person Map.
Instrument Quality

Unidimensionality. This aspect was measured by the Rasch Model with residual principle component analysis (PCA) of measuring the uniformity of the instrument to determine the measurement (Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). The measurement results displayed a raw variance of 30.6% and 12% variances that cannot be explained by the instrument. This means that the minimum unidimensionality requirements of 20% and variance that cannot be explained by the instrument ≥15% are fulfilled (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015).

Rating Scale. This is a test to verify the ranking options used in the SKKM instrument that may confuse participants (Bond & Fox, 2015; Dimitrov, 2012; Linacre, 2019; Sumintono & Widhiarso, 2014; 2015). The results of the Rasch Model analysis revealed that the average observation value starts from logit -2.67 for the choice of score 1 (very irrelevant), logit -1.17 for the choice of score 2 (irrelevant), logit -0.08 for the choice of score 3 (quite relevant), logit +1.13 for choice of score 4 (relevant) and logit +2.87 for choice of score 5 (very relevant). Similar results are also displayed by Andrich Threshold, which tested the accuracy of the politemical value. This value shows the movement from NONE to negative (-1.27, -.68) and continues to lead to positive (.37, 1.58) in sequence. To be exact, the choice of ranking scale used is valid and appropriate for participants.
Reliability. Table 1: Summary of Statistics shows person measure = +.36 logit. A higher average value of logit .0 indicates the tendency of participants to answer in each item. Cronbach’s Alpha value for measuring test reliability is the interaction between person and item as a whole shows .82, which is good. Meanwhile, the value of person reliability is .79 and item reliability .99. Therefore, it can be concluded that the consistency of participant answers was quite good (Sumintono & Widhiarso, 2014; 2015). Infit and Outfit Mnsqs display similar results for person and item. The average values are sequentially 1.02 and 1.01 and 1.00 and 1.01, which are close to the ideal value of 1.00. Likewise, Infit and Zstd Outfit presents the average value of the person and items as -.5 and -.5 and -.2 and .0, which means good because it is close to .0.

The separation value is used to group people and items. The greater the separation value indicates the better the quality of the instrument as it can identify groups of respondents and item groups. The equation used was $H = \frac{[(4 \times 1.94) + 1]}{3} = 2.92$. The result was rounded up to 3. This result means that there are three groups of respondents.

<table>
<thead>
<tr>
<th>Output</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Item reliability</td>
</tr>
<tr>
<td>Separation index</td>
<td>8.41</td>
</tr>
<tr>
<td>Separation strata (H)</td>
<td>2.92</td>
</tr>
<tr>
<td>Higher logit value</td>
<td>+ 1.10 logit (X_{35})</td>
</tr>
<tr>
<td>Lower logit value</td>
<td>-1.09 logit (X_{38})</td>
</tr>
<tr>
<td>Person</td>
<td>Person reliability</td>
</tr>
<tr>
<td>Separation index</td>
<td>1.94</td>
</tr>
<tr>
<td>Separation strata (H)</td>
<td>2.92</td>
</tr>
<tr>
<td>Higher logit value</td>
<td>+ 1.82 logit (8LDPI)</td>
</tr>
<tr>
<td>Lower logit value</td>
<td>-.40 logit (0PDID)</td>
</tr>
<tr>
<td>Instrument</td>
<td>Cronbach’s alpha</td>
</tr>
<tr>
<td>Raw variance explained by measures</td>
<td>30.6</td>
</tr>
<tr>
<td>Unexplained variance in 1st contrast</td>
<td>12.0</td>
</tr>
<tr>
<td>Unexplained variance in 2nd contrast</td>
<td>9.6</td>
</tr>
</tbody>
</table>

Table 1. Summary Statistics
Conclusion
This research has produced a measure of the wisdom of multicultural counselor’s candidates with satisfying psychometric properties. The New MC-WiSe is proven to reveal one dimension of the wisdom of multicultural counselors. The New MC-WiSe has satisfied psychometric properties based on the criteria of item measure, order fit items, DIF, unidimensionality, person measure, order person fit, item-person maps, rating scale, test reliability, person reliability and item reliability. The final MC-WiSe numbered 22 items to measure 2 dimensions and 11 indicators. There are 12 items to measure cognitive dimensions and 10 items to measure affective and awareness dimensions. Two items represent each indicator. The study also exposed the structure of the factors are still needed to be improved. Further studies are needed to develop and to test instruments with confirmatory factor analysis in broader participants, both prospective counselors and counselors with regard to their socio-demographic proportions.

Acknowledgements
We sincere appreciation is conveyed to the research participants and those who helped to publish this article.
REFERENCES


36
