© 2024 The UNM Mentoring Institute
The Chronicle of Mentoring & Coaching,
2024, Vol. 8, No. 2
doi:10.62935/iq66350

The Chronicle of Mentoring & Coaching

Measuring Faculty Mentoring Competency: Establishing the Validity of a Short Form

Shujin Zhong[1] & Nicole Maccalla[2]
[1]University of North Florida
[2]University of California—Los Angeles

This paper describes the procedures for evaluating the psychometric properties of the 26-item Mentoring Competency Assessment (MCA) scale and developing short-forms to measure faculty mentoring outcomes for the NIH-funded Building Infrastructure Leading to Diversity (BUILD) Initiative and National Research Mentoring Network (NRMN). Analyses were conducted using responses to the MCA scale from NRMN mentors and faculty across 11 BUILD institutions in the 2017-18 academic year. After performing extensive item factor analyses and taking the MCA sub-constructs into analytical consideration, we created an 8-item short form and a 14-item short form. Analyses indicate that both short forms nearly equivalently measure faculty mentoring competency and are more feasible to implement in future studies, compared to the original, longer scale.

Keywords: Mentoring Competency Assessment scale, faculty mentoring competence, Building Infrastructure Leading to Diversity, National Research Mentoring Network

Acknowledgements: We would like to thank members of the DPC for their steadfast contributions and all Enhance Diversity Study participants for their willingness to contribute. The authors prepared this manuscript on behalf of the Diversity Program Consortium. Work reported in this publication was supported by the Office of The Director, National Institutes of Health Common Fund and Office of Scientific Workforce Diversity awards UL1GM118979, UL1GM118976, UL1GM118973, UL1GM118964, UL1GM118985, UL1GM118991, UL1GM118982, UL1GM118988, UL1GM118970, UL1GM118967, U54GM119024, U24GM132176, and U24GM132217 administered by the National Institute of General Medical Sciences. The work is solely the responsibility of the authors and does not necessarily represent the official view of the National Institutes of Health.

Introduction

The US is becoming more racially and ethnically diverse, and this trend is expected to grow in the coming decades (Cohn & Caumont, 2016). In sharp contrast to this trend is the demography of the biomedical research workforce, which lacks diversity despite numerous efforts to increase diversity for the past 40 years (Valantine & Collins, 2015; McGee, Saran, & Krulwich, 2012). The disparity in diversity between the general population and that of biomedical researchers can be reduced if the candidate pool of the biomedical research field is a diverse population. However, the completion of science-related college degrees of underrepresented groups (URGs, as defined in Maccalla et al., 2021) has continuously fallen behind (Rask, 2011; Valantine & Collins, 2015).

In response to these continued disparities in the biomedical workforce, the National Institutes of Health (NIH) announced funding for a new set of initiatives in 2013. The aim was unique in that it employed a broad approach to "enhance diversity in the NIH-funded biomedical, behavioral, clinical, and social sciences research workforce" through innovative and transformative approaches to "undergraduate research training and mentorship"(Funding Opportunity Announcement, 2013, p. 1) The funding opportunity resulted in the creation of "Enhancing the Diversity of the NIH-Funded Workforce, also known as the Diversity Program Consortium (DPC), a trans-NIH program" managed by the National Institute of General Medical Sciences (NIGMS, 2024).

The first phase of the DPC consisted of three core integrated initiatives: the Building Infrastructure Leading to Diversity (BUILD) Initiative, the National Research Mentoring Network (NRMN), and the Coordination and Evaluation Center (CEC). Since its launch in 2014, this national collaboration has worked to advance the overarching goal of developing, implementing, assessing, and disseminating innovative and effective approaches to engaging, training, and mentoring students, enhancing faculty development, and strengthening institutional research training infrastructure. Undergraduate student participants in BUILD receive training and mentoring during their undergraduate studies from BUILD-trained faculty mentors. By providing holistic training and mentoring for faculty and students, the program contributes to "engaging and retaining students from diverse backgrounds in biomedical research" (NIGMS, 2024). The NRMN is developing a national network of mentors and mentees from a range of

biomedical fields and career stages to "advance careers of individuals from diverse backgrounds" (NIGMS, 2024).The Enhance Diversity Study (EDS), led by the CEC, assessing how individual BUILD programs are achieving desired outcomes (aka Hallmarks of Success) through a multi-site, national, longitudinal, equity-focused impact evaluation (Davidson et al., 2017; Guerrero et al., 2022; McCreath et al., 2017). The EDS includes annual surveys to students and faculty at 10 BUILD programs, located in 11 institutions from 2015 to 2024 (Davidson et al., 2017; McCreath et al., 2017). Preparation for and mentoring of undergraduate and graduate students, post-docs, and junior faculty is a key focus for the DPC activities. It is expected that students participate in mentored research, receive frequent mentoring, and are satisfied with the quality of mentorship. Faculty are expected to engage in mentor training and mentoring frequently and do so with high self-efficacy while utilizing evidence-based practices (McCreath et al., 2024).

In Homer's Odyssey, a mentor was a "wise and trusted counselor" whom Odysseus entrusted with the care and education of his son, Telemachus. Along these lines, Crisp and Cruz (2009) defined a mentor as "a wise, responsible, and trusted advisor," who guides the development of an individual (p. 527). Previous studies have concluded that mentorship programs are beneficial to the outcomes of individuals and to the development of organizations or institutions (Byars-Winston et al., 2015; Eagan et al., 2013; Lin & Hsu, 2012; McCreath et al., 2024; Seibert, 1999). The faculty-student mentoring model is the most common form of mentorship in college (Zhong, 2016), and the effectiveness of faculty mentoring has been considered one of the important indicators of institutional change (Lin, & Hsu, 2012; Seibert, 1999). However, our understanding of what specific aspects of mentoring lead to students' career advancement and improved psychosocial support is largely anecdotal (Crisp & Cruz, 2009; Gómez, Ali, & Casillas, 2014). Jacobi (1991) and Crisp and Cruz (2009) holistically reviewed studies on mentorship in higher education and note limitations including small sample sizes, a lack of evidence from quasi-experimental designs, and measurement issues that led to low external validity.

In 2010, five awardee institutions of the National Institutes of Health (NIH) Clinical and Translational Science Awards (CTSA) jointly developed a clinical and translational research mentor training curriculum, based on the Entering Mentoring seminar, a mentor training program in the science, technology, engineering, and mathematics (STEM) fields (Pfund et al., 2013). This training program, featured a large quasi-experimental nationwide study that involved participants from 16 U.S. universities, remains a remarkable component of the mentorship literature, overcoming many of the previously mentioned study limitations. This experimental program had a fairly large sample size of 283 pairs of mentors and mentees, and generated evidence of high external validity.

In evaluating this program, Fleming et al. (2013) utilized the Mentoring Competency Assessment (MCA) scale for assessing mentor training and fulfilling the requirement of the CTSA (Pfund et al., 2013). The MCA is a 26-item scale that measures six sub-domains of mentoring competency (Appendix 1): maintaining effective communication, aligning expectations, assessing understanding, addressing diversity, fostering independence, and promoting professional development (Fleming et al., 2013). The original MCA scale went through a rigorous development process, including 1) reviewing of the Mentorship Effectiveness Scale (by Berk et al., 2005) and other instruments, 2) interviews with mentoring program participants, and 3) scale validation procedures on 283 pairs of mentors (self-rating) and mentees (rating their mentors) at 16 U.S. universities using confirmatory factor analysis (CFA).

The articles by Pfund et al. (2013) and Fleming et al. (2013) inspired later programs and evaluations of those programs, including the use of the MCA to study the effects of mentor training on mentorship. The MCA was used in the national longitudinal EDS as part of the evaluation of the DPCs BUILD and NRMN Phase I initiatives, supported by the National Institutes of Health (Davidson, et. al, 2017, Guerrero et. al., 2022). Although the MCA provides a good item bank for measuring mentoring, it proved difficult to administer in its entirety due to the length of the measure. When juggling the burden on respondents and the maintenance of high response rates, it becomes vital for researchers to maximize efficiency and scale validity. This paper examines the generalizability and transferability of the MCA scale in a large national training program and explores variations of the MCA scale's appearance in large-scale surveys in higher education in an effort to create a shorter, more feasible form to implement in the context of a longitudinal study.

Methods

Using survey data to evaluate program effectiveness hinges upon having valid and reliable measurement tools. A common approach is to identify existing valid measures of the construct of interests from the literature. In this process, we must ensure that the identified scales and items are measuring similar populations between the original scale development work and subsequent studies.

Although the MCA provides a good item bank for measuring mentor competencies, to fully employ the items in the EDS surveys, a localization and refinement process of tailoring the scale to fit the BUILD population was necessary for several reasons. First, the MCA was designed to measure mentoring in a specific program for clinical and translational researchers, instead of faculty mentoring in general. Faculty mentoring

in the BUILD program focuses on mentoring students, especially undergraduate students, in classroom and research settings. Second, the MCA was developed based on pairs of mentors and mentees, and the item selection considered results both from mentors' self-rating and mentees' rating of mentors. In the EDS, we focus on mentors' self-ratings. Third, measuring faculty mentoring only takes a small portion of the BUILD and NRMN evaluations, and a 26-item scale was seen as redundant for a single construct and burdensome for respondents. It is possible that the construct was "overly measured" by too many items. More items will always provide more information; however, in survey administration, there is a debate between collecting adequate information and survey length – if the survey is too long, it might jeopardize the completion rate. Therefore, after identifying the most informative items in the scale, a further step is to tailor the scale to a shorter form for a better completion rate. Figure 1 presents the theoretical framework of this study. The ultimate goal is to find a short unidimensional scale measuring faculty mentors' self-assessed mentor competency for the DPC population.

In this paper, our goal is to assess the psychometric properties of the MCA and increase the feasibility of using the MCA scale by developing short forms to measure mentoring competency in comprehensive, outcome-focused surveys. One feature of many large-scale studies, including the EDS, is the measurement of multiple outcomes of interest. When scales include a multitude of items, surveys can quickly become unwieldy in length, and pursuing item reduction procedures becomes necessary. From a practical consideration, since survey length and time to completion are negatively correlated with response rates and survey completion, having the fewest number of items on surveys is of paramount importance. Item reduction is built upon getting the "correct" measurement model, with reduced scales maintaining the overall integrity of the original scale. This paper details the results from item reduction procedures applied to the original form (26 items) of the Mentoring Competency Assessment (Fleming et al., 2013). The authors coordinated this expanded effort with the original developers of the scale.

To establish that the scales could adequately measure mentor competency in the DPC population, a scale validation procedure is necessary. This is important because validity is not an inherent property of a survey instrument and is primarily determined by the specific purpose of data collection with particular settings and participants.

Enhance Diversity Study investigators set out to answer the following research questions:

1. To what extent can the MCA be shortened from its original 26-item length, while maintaining similar psychometric properties?
2. Which items across the six MCA sub-domains perform best?
3. How well are faculty mentorship skills measured by the short form in comparison to the original MCA scale?

The purpose of this study is to establish a short form of the Mentoring Competency Assessment (MCA-short) while maintaining similar psychometric properties of the original scale. Sequential statistical analyses employed to test the quality of each item and the overall scale include Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), Item Response Theory (IRT), and cross-validation procedures (mentor/mentee ratings). It is important that the high validity and reliability of the original MCA scale transfer to the MCA-short. Pairing down the assessment of mentoring skills to a reasonable length on the EDS surveys supports complete and sustained engagement in the longitudinal study of DPC effectiveness. The MCA-short is being used in subsequent EDS faculty/mentor follow-up surveys and studies associated with the DPC.

The data utilized in this paper are faculty/mentor responses to the original MCA scale from the BUILD Faculty Annual Follow-Up Survey (FAFS 2017-18) and the 2018 NRMN Mentor Survey. The BUILD faculty are from the 10 BUILD grantees and 11 corresponding BUILD sites. The FAFS was distributed to faculty members in the BUILD sites and received 683 responses. Among them, 586 participants self-identified as mentors, and 564 participants answered at least one question on the MCA scale. The 2018 NRMN Mentor Survey was distributed to NRMN mentors and received 1107 responses. Among those 754 who self-identified as mentor, 621 participants answered at least one question on the MCA scale.

In this study, we follow a similar procedure as the one implemented by the MCA scale developer to validate a short form of the MCA using data from mentors' self-ratings. Primarily, we assessed faculty/mentor self-ratings and performed exploratory item-factor analysis to determine the quality of the items with the FAFS survey data. As mentioned, the instrument developers suggested that the 26 items could be decomposed into 6 domains (Fleming, et al., 2013). Factor analysis helps identify item clusters and construct measurement structures (i.e., testing the dimensionality). Since we are assessing the responses to the MCA scale, the dimensionality proposed by Fleming, et al. (2013) is considered.

Dimensionality analysis is a good first step when data are collected for assessing ability or competency. If the scale is not unidimensional, it is difficult to explain which it. In the original MCA scale (see Appendix 1), all the items are coded in the same direction, with a smaller number representing higher frequency/rating and a larger number representing lower frequency/rating. The item responses are all Likert-type ordinal

responses, which requires the use of polychoric correlations to generate EFA results. After obtaining polychoric correlations, multiple simple statistical rules of thumb, such as scree plots and parallel analysis provide a sense of dimensionality. The goal of exploration is to find a model that has the most suitable number of factors aligned with the mentoring outcomes of interest, and at the same time, fit a structure that is easy to interpret.

Hansen et al. (2014) used item-explained common variance (I-ECV) to assess whether an item is strongly associated with the general dimension. With an I-ECV above .5, the item is strongly related to the general dimension; with an I-ECV above .8, the item is weakly influenced by a group-specific dimension. These criteria are considered as our guide when determining the final item selection in our study.

After generating the results from the exploratory analyses of the FAFS faculty/mentor data, we conducted confirmatory factor analyses on the NRMN mentor data to check if the determined dimensionality appeared to be the same and if the selected items also showed high quality with a second DPC population.

Another framework that is often used for scale development and item analysis is Item response theory (IRT). IRT is especially helpful with obtaining test scoring and assessing the growth of latent traits (Thissen, & Steinberg, 2009; Thissen, 2016). IRT models are logistic models that can be used to model the probability of selecting a certain response to an item, using ability and item parameters. The difficulty parameter describes the location where the amount of the latent trait has a .5 probability of endorsing the item. There can be multiple locations if an item has polytomous responses. The discrimination parameter is the slope of the curve at the item location. In an item with ordered polytomous categories, the discrimination parameter is the same across the ordinal responses under the graded response model (as well as the partial credit model). From the parameters, researchers can plot Item Characteristic Curves (ICC), Item Information curves (IIC), and Test Information Curves (TIF). These curves can help with examining item qualities for the given population, identifying collapsed responses, and visualizing the extent to which the constructs are measured.

We use Fisher Information to determine how much information an item can provide at each point and to what extent the item can help with the classification of individuals' latent scores. To generate the Fisher Information, we first fit the data into the graded response model and established the discrimination parameters and difficulty parameters. In an item with ordered polytomous categories, the discrimination parameter is the same across the ordinal responses, under the graded response model. To compare if the created short forms could yield a similar conclusion, we used Fisher scoring to obtain individual IRT scores and checked the correlation and rank order correlation among the scores. The goal of item selection for creating a short form is to group the most informative and unbiased items that can reflect the same item structures and response structures of the original scale, while measuring the overall construct adequately well for utilization in longitudinal analyses.

Results

The results from the polychoric correlation, using FAFS data are displayed in Appendix 2. All items are positively and highly correlated, a sign of scale unidimensionality. In the exploratory analyses, a 1-factor model (Appendix 3) with a 1-factor solution is deemed sufficient with high communalities (shared variances). Parallel analysis (Appendix 4), based on a polychoric correlation matrix and reduced matrix (SMCs placed on the diagonal), both support the 1-factor model. To further examine the dimensionality, exploratory factor models with different numbers of factors were tested, presenting the 4-factor, 5-factor, and 6-factor solutions under oblique rotation (Appendix 5). The six sub-domains, unlike those described in Fleming et al., do not stand out clearly. Several items load relatively high on multiple domains, a sign that the 6-factor model is over-extracted. The correlations between factors are too high to indicate separate domains. To further examine if a 6-factor model can be a good fit, the item structure from Fleming et al. (2013) is used to form the target rotation matrix (see Appendix 1). Appendix 6 reports the 6-factor model under target rotation, still suggesting over-extraction, with highly correlated factors.

Bi-factor models were tested , where all 26 items fall under one general mentoring domain, but also allowing some items to belong to a specific sub-domain. The bi-factor models with five sub-domains and six sub-domains (Appendix 7) indicate that a bi-factor model could be a better fit. Item level Expected Common Variances (I-ECV, the variances that loaded on the main faculty mentoring domain) are high in general, supporting the unidimensional assumption. The sub-domains are all orthogonal. Similar to the previous procedure, a target rotation matrix was produced, leading to the bi-factor 6 sub-domain results and its corresponding residual matrix (Appendix 8). The comparisons among CFA results are presented in Appendix 9. CFA goodness-of-fit indexes from OLS and from the Maximum Likelihood (ML) estimation (since rotation methods would not influence the values) are reported. Also reported are the degrees of freedom (df), discrepancy function (F_{min}), Chi-square ($\chi 2$), CFI, TLI, NNFI, RMSEA and its 90% confidence interval, RMR, SRMR, and ECVI for reference. From the goodness-of-fit indexes, we concluded that the bi-factor model should be selected, which somewhat reflected, but still yielded a different result from the original analytical result by Fleming et al. (2013).

From EFA, two elimination criteria should be

considered: one is the I-ECV, which indicates the level of unidimensionality, and the other is the residual matrix, which reflects the accuracy of communality estimation. Ideally, items with high I-ECV and low residuals should remain. In this report, the preliminary cut-offs are placed at an I-ECV below .5 and residuals above .05. Under the condition that we still hope to keep the 6 sub-domains proposed by Fleming et al., one additional criterion for consideration is keeping at least two items in each sub-domain, which also indicates that the selected items will at most have a high loading on one sub-domain in addition to the general mentoring factor. The 8 items in Appendix 10 are suggested for inclusion in a Faculty MCA short form. In Appendix 11, we present other short forms without considering the Fleming et al. (2013) sub-constructs, which are not considered the ideal solution, as the sub-domains remain supported by literature.

Using the NRMN mentor data, we conducted a confirmatory factor analysis to check if the determined dimensionality appeared to be the same in a similar DPC population and if the selected items exhibited good quality. We, again, use I-ECV to assess item quality. Almost all selected items have I-ECV over .7, and the explained variances are, in general, higher than those from the EFA. The bi-factor model best fits the data , with the CFI over .94 and the TLI over .93, showing an adequate model fit.

For IRT analyses, we fit a graded response model to the FASF data to estimate the discrimination and difficulty parameters. From the parameters, Test Information Curves (TIC) are plotted. These curves can help with examining item qualities for the given population, identifying collapsed responses, and visualizing the extent to which the constructs are measured. For our population, most items perform well between -3SE to +2SE. The comparison of TICs (Appendix 12) for the original scales and short forms indicates that either short form can provide sufficient information, with the 8-item short form functioning best.

Through IRT analysis, we also calculate the individual scores of faculty mentoring. When comparing the scores from the 8-item, 14-item, and 26-item scales, we find that the correlations of the scores are relatively high. The 8-item scores are .960 correlated with the original MCA, and the rank-order correlation is .953; the 14-item scores are .982 correlated with the original MCA, and the rank-order correlation is .978. With this additional information, we conclude that the 8-item and the 14-item short forms are good substitutes for the original MCA scale of measuring faculty mentoring outcomes. The 8-item scale is suggested for use with the DPC faculty population.

Discussion

In this paper, we examined the psychometric properties of the original MCA using faculty/mentor survey data from 2017-2018. Sequential statistical analyses employed to test the quality of each item and the overall scale included Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), and Item Response Theory (IRT). The resulting short forms include an 8-item scale (MCA-short) and a 14-item scale across the original 6 sub-domains. The MCA-short achieved high reliability (over .92) in the DPC faculty/mentor population, with most items performing well between -3SE to +2SE. When comparing the IRT scores from the 8-item and the 26-item scales, we find that the correlations of the scores are relatively high. We can conclude that the 8-item short form and the 14-item short form are good substitutes for the original MCA scale of measuring faculty mentoring for the DPC faculty/mentor population. Given these results, the 8-item short form was re-named “MCA-short” and has been used with subsequent Enhance Diversity faculty/mentor follow-up surveys and studies associated with the Diversity Program Consortium. One great benefit for the Enhance Diversity Study is the survey space saving implications and potential impacts on minimizing participant burden and increasing response/completion rates.

The MCA-short (Table 10) is an 8-item short form of the original MCA. The MCA-short proves itself to be both valid and reliable and is offered as an alternative to the long form in measuring faculty mentoring competency, particularly when researchers are concerned about space constraints and/or respondent burden. The Mentoring Competency Assessments (MCA and MCA-short) have been used to evaluate evidence-based mentoring skills from the perspective of faculty/mentors in the Enhance Diversity Study. Permission from original MCA authors has been granted to publish the MCA-short.

The short form of the mentoring competency assessment (MCA-short) maintains the general measurement properties and functionality of the original MCA, with all six sub-domains. Shorter forms may be more feasible to implement in future studies where tracking mentoring outcomes is important. This study demonstrates the refinement of an existing scale, maintaining content and construct validity while yielding a practical instrument for collecting credible evidence. Findings support the use of the MCA-short in biomedical training programs aimed at improving faculty mentorship skills. Scores from the MCA-short can be used to gauge faculty mentoring skills, identify areas for mentor development, assess convergence/divergence between mentor and mentee perspectives, and track changes over time. The MCA-short form was used with subsequent EDS faculty/mentor follow-up surveys and related outcome-focused studies associated with the DPC. Given prior research on the significance of quality mentorship in supporting the academic and career progression of students from underrepresented backgrounds, quality measurement tools, such as the MCA-short, are of great importance to the field. Future studies such as other variations of the MCA-

short (e.g., Zhong, 2021) have been performed using similar approaches as demonstrated in this study.

References

Berk, R. A., Berg, J., Mortimer, R., Walton-Moss, B., & Yeo, T. P. (2005). Measuring the effectiveness of faculty mentoring relationships. *Academic Medicine, 80*(1), 66-71.

Byars-Winston, A., Branchaw, J., Pfund, C., Leverett, P., & Newton, J. (2015). Culturally diverse undergraduate researchers' academic outcomes and perceptions of their research mentoring relationships. *International Journal of Science Education, 37*(15), 2533–2554. https://doi.org/10.1080/09500693.2015.1085133

Cohn, D. V., & Caumont, A. (2016). *10 demographic trends shaping the US and the world in 2016.* Pew Research Center.

Crisp, G., & Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education, 50*(6), 525-545

Davidson, P.L., Maccalla, N.M.G., Afifi, A.A., Guerrero, L., Nakazono, T., Zhong, S., & Wallace, S.P. (2017). A participatory approach to evaluating a national training and institutional change initiative: the BUILD longitudinal evaluation. *BMC Proceedings 2017, 11*(Suppl 12):15.

Eagan Jr, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a difference in science education: The impact of undergraduate research programs. *American educational research journal, 50*(4), 683-713.

Fleming, M., House, M. S., Shewakramani, M. V., Yu, L., Garbutt, J., McGee, R., ... & Rubio, D. M. (2013). The mentoring competency assessment: validation of a new instrument to evaluate skills of research mentors. *Academic Medicine: Journal of the Association of American Medical Colleges, 88*(7), 1002.

Funding Opportunity Announcement. (2013). RFA-RM-13-016: NIH Building Infrastructure Leading to Diversity (BUILD) Initiative (U54). National Institutes of Health (NIH). https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-13-016.html

Gómez, R. L., Ali, A., & Casillas, W. (2014). Mentorship and the professional development of culturally responsive evaluators in the American evaluation association's graduate education diversity internship (GEDI) program. *New Directions for Evaluation, 2014*(143), 49-66.

Guerrero L.R., Seeman T.E., McCreath H.E., Maccalla N.M.G., Norris K.C. (2022). Understanding the context and appreciating the complexity of evaluating the Diversity Program Consortium. *New Directions for Evaluation, 174,* 11-20. doi: 10.1002/ev.20506.

Hansen, M., Cai, L., Stucky, B. D., Tucker, J. S., Shadel, W. G., & Edelen, M. O. (2014). Methodology for developing and evaluating the PROMIS® smoking item banks. nicotine & tobacco research, 16(Suppl_3), S175-S189.

Jacobi, M. (1991). Mentoring and undergraduate academic success: A literature review. *Review of educational research, 61*(4), 505-532.

Lin, Y. N., & Hsu, A. Y. P. (2012). Peer mentoring among doctoral students of science and engineering in Taiwan. *Asia Pacific Education Review, 13*(4), 563-572.

Maccalla, N. M. G., Gutierrez, A., Zhong, S., Wallace, S. P., & McCreath, H. E. (2020). Technical report: Evaluation of post-secondary student outcomes: Underrepresented (URG) and well-represented (WRG) group variable construction in the enhance diversity study using the November 2019 nih guidelines.

McCreath, H. E., Norris, K. C., Calderón, N. E., Purnell, D. L., Maccalla, N. M., & Seeman, T. E. (2017, December). Evaluating efforts to diversify the biomedical workforce: the role and function of the Coordination and Evaluation Center of the Diversity Program Consortium. *BMC Proceedings, 11*(12), 27. BioMed Central.

McCreath, H., Eagan, K.M., Maccalla, N.M.G., Joseph, C.J., & Norris, K.C. (2024). BUILDing Engaged Mentors: Examining the Efficacy of BUILD-led Mentor Training. *The Chronicle of Mentoring and Coaching.* https://doi.org/10.62935/wm3295 In Pfund, C., Sorkness, C. A., Trube, M.B., Myers, O. B., Dominguez, N., & Martinez, W. (Eds.). (2024). Mentorship interventions across career stages in biomedical & health sciences fields. [Special Issue], *The Chronicle of Mentoring & Coaching, 8*(1). https://doi.org/10.62935/wn7125

McGee Jr, R., Saran, S., & Krulwich, T. A. (2012). Diversity in the biomedical research workforce: developing talent. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine, 79*(3), 397-411.

National Institute of General Medical Sciences (NIGMS). (2024, June 6). Enhancing the Diversity of the NIH-Funded Workforce. NIH and NIGMS.https://www.nigms.nih.gov/training/dpc

Pfund, C., House, S., Spencer, K., Asquith, P., Carney, P., Masters, K. S., ... & Fleming, M. (2013). A research mentor training curriculum for clinical and translational researchers. *Clinical and Translational Science, 6*(1), 26-33.

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of*

Education Review, 29(6), 892-900.
Seibert, S. (1999). The effectiveness of facilitated mentoring: A longitudinal quasi-experiment. *Journal of Vocational Behavior, 54*(3), 483-502.
Thissen, D. (2016). Bad questions: An essay involving item response theory. Journal of *Educational and Behavioral Statistics, 41*(1), 81-89.
Thissen, D., & Steinberg, L. (2009). Item response theory. *The Sage handbook of quantitative methods in psychology* (148-177).
Valantine, H. A., & Collins, F. S. (2015). National Institutes of Health addresses the science of diversity. *Proceedings of the National Academy of Sciences, 112*(40), 12240-12242.
Zhong, S. (2016). *The effectiveness of a peer mentorship program: A mixed methods study.* University of Southern California.
Zhong, S. (2021). *Improving measurements in large-scale surveys and using survey data to assess program impacts.* University of California, Los Angeles.

About the Authors

Shujin Zhong, PhD, is an Assistant Professor at the University of North Florida. Her research focuses on educational measurement, latent variable modeling, causal inference, social network analysis, research design, and the application of methods in the field of higher education and STEM education. Before joining the University of North Florida, she was a Postdoctoral Associate in the Department of Human Development and Quantitative Methodology at the University of Maryland. Her prior work, during doctoral training, was with the NIH Diversity Program Consortium (DPC), Coordination and Evaluation Center (CEC) at UCLA.

Nicole Maccalla, PhD, is an Academic Administrator in the Department of Education at the University of California, Los Angeles, and serves as the Co-Lead of the CEC Evaluation Core, CEC editor-in-chief, and chair of the Diversity Program Consortium Evaluation Implementation Working Group. Dr. Maccalla's research interests include underrepresentation in higher education, achieving diversity, equity, and inclusion in organizations, evaluation for organizational capacity building, meaningful measures of teaching effectiveness, and scale development.

Appendix

Appendix I
The 26-item Mentoring Competency Scale

Sub-domains	Skills	Items
Communication	Listening	Active listening
	Feedback	Providing constructive feedback
	Trust	Establishing a relationship based on trust
	Styles	Identifying and accommodating different communication styles
	Strategies (C)	Employing strategies to improve communication with mentees
	Coordinate	Coordinating effectively with mentees' other mentor
Expectation	Set expectations	Working with mentees to set clear expectations of the mentoring relationship
	Align expectations	Aligning expectations with mentees'
	Differences	Considering how personal and professional differences may impact expectations
	Goals	Working with mentees to set research goals
	Strategies (E)	Helping mentees to develop strategies to meet goals
Assessing	Knowledge	Accurately estimating mentees' level of scientific knowledge
	Mentee ability	Accurately estimating mentees' ability to conduct research
	Mentee skills	Employing strategies to enhance mentees' knowledge and abilities
Independence	Motivation	Motivating mentees
	Confidence	Building mentees' confidence
	Creativity	Stimulating mentees' creativity
	Contributions	Acknowledging mentees' professional contributions
	Negotiating	Negotiating a path to professional independence with mentees
Diversity	Prejudice	Taking into account the biases and prejudices the mentor brings to the mentor/mentee relationship
	Background	Working effectively with mentees whose personal background is different from the mentor (age, race, gender, class, region, culture, religion, family composition etc.)
Profession	Network	Helping mentees network effectively
	Career goals	Helping mentees set career goals
	Work/life balance	Helping mentees balance work with their personal life
	Role model	Understanding the mentor's impact as a role model
	Acquire resources	Helping mentees acquire resources (e.g. grants, etc.)

Rating on a 7-point Likert-type scale: 1 = not at all skilled, 4 = moderately skilled, 7 = extremely skilled

Appendix II
Correlation Matrix

	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22	I23	I24	I25	I26
I1																										
I2	0.72																									
I3	0.72	0.7																								
I4	0.59	0.64	0.66																							
I5	0.65	0.65	0.65	0.83																						
I6	0.45	0.52	0.45	0.53	0.58																					
I7	0.52	0.58	0.52	0.55	0.63	0.67																				
I8	0.54	0.64	0.56	0.65	0.66	0.65	0.82																			
I9	0.57	0.55	0.56	0.62	0.63	0.51	0.58	0.71																		
I10	0.46	0.62	0.49	0.51	0.54	0.43	0.56	0.65	0.58																	
I11	0.6	0.61	0.53	0.61	0.65	0.56	0.67	0.69	0.63	0.65																
I12	0.38	0.54	0.43	0.47	0.48	0.46	0.48	0.58	0.45	0.64	0.6															
I13	0.45	0.58	0.45	0.47	0.49	0.45	0.52	0.58	0.47	0.69	0.62	0.87														
I14	0.55	0.63	0.55	0.64	0.71	0.54	0.59	0.65	0.66	0.69	0.75	0.66	0.65													
I15	0.47	0.53	0.52	0.57	0.62	0.48	0.57	0.6	0.62	0.57	0.66	0.47	0.51	0.78												
I16	0.56	0.57	0.61	0.61	0.61	0.46	0.53	0.64	0.58	0.56	0.71	0.51	0.51	0.71	0.71											
I17	0.55	0.61	0.56	0.57	0.58	0.52	0.57	0.62	0.58	0.55	0.63	0.51	0.52	0.68	0.67	0.74										
I18	0.49	0.62	0.52	0.47	0.52	0.45	0.46	0.56	0.56	0.66	0.6	0.5	0.52	0.63	0.61	0.67	0.63									
I19	0.52	0.61	0.57	0.6	0.62	0.56	0.59	0.66	0.55	0.62	0.7	0.6	0.62	0.7	0.66	0.7	0.69	0.71								
I20	0.6	0.54	0.59	0.64	0.68	0.5	0.55	0.59	0.72	0.51	0.65	0.46	0.47	0.65	0.65	0.62	0.64	0.53	0.61							
I21	0.59	0.53	0.56	0.59	0.58	0.41	0.46	0.49	0.63	0.45	0.59	0.39	0.44	0.57	0.56	0.61	0.57	0.51	0.51	0.73						
I22	0.43	0.45	0.55	0.5	0.54	0.47	0.52	0.53	0.5	0.39	0.53	0.37	0.37	0.55	0.56	0.59	0.58	0.5	0.63	0.53	0.58					
I23	0.55	0.59	0.59	0.56	0.6	0.49	0.58	0.61	0.62	0.61	0.7	0.51	0.54	0.68	0.69	0.67	0.66	0.7	0.74	0.64	0.64	0.69				
I24	0.43	0.44	0.47	0.56	0.56	0.47	0.43	0.51	0.53	0.43	0.55	0.39	0.43	0.55	0.59	0.6	0.55	0.46	0.56	0.61	0.6	0.57	0.65			
I25	0.54	0.55	0.62	0.63	0.69	0.59	0.57	0.64	0.64	0.53	0.62	0.44	0.47	0.65	0.68	0.71	0.68	0.6	0.68	0.71	0.64	0.66	0.74	0.69		
I26	0.47	0.45	0.48	0.46	0.5	0.42	0.47	0.44	0.44	0.5	0.53	0.38	0.38	0.54	0.55	0.54	0.58	0.51	0.58	0.5	0.46	0.62	0.65	0.57	0.64	

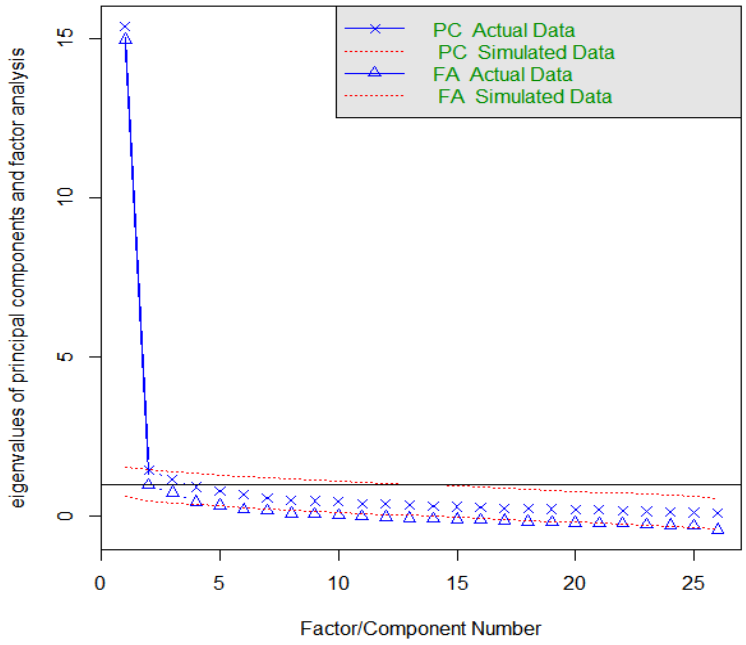
*Blue represents higher values, and red lower values
(the numbers are rounded due to the space limit)

Breakdowns follow: 1) maintaining effective communication, 2) aligning expectations, 3) assessing understanding, 4) fostering independence, 5) addressing diversity, and 6) promoting professional development.

Appendix III
1 Factor Solution

Item	Factor Loading	Communality	Uniqueness
Active listening	0.707	0.500	0.500
Providing constructive feedback	0.761	0.579	0.421
Establishing a relationship based on trust	0.732	0.536	0.464
Identifying and accommodating different communication styles	0.768	0.590	0.410
Employing strategies to improve communication with mentees	0.808	0.653	0.347
Coordinating effectively with your mentees' other mentor	0.661	0.437	0.563
Working with mentees to set clear expectations of the mentoring relationship	0.738	0.544	0.456
Aligning your expectations with your mentees'	0.809	0.654	0.346
Considering how personal and professional differences may impact expectations	0.765	0.586	0.414
Working with mentees to set research goals	0.731	0.535	0.465
Helping mentees to develop strategies to meet goals	0.828	0.685	0.315
Accurately estimating your mentees' level of scientific knowledge	0.658	0.433	0.567
Accurately estimating your mentees' ability to conduct research	0.685	0.469	0.531
Employing strategies to enhance your mentees' knowledge and abilities	0.847	0.718	0.282
Motivating your mentees	0.789	0.623	0.377
Building mentees' confidence	0.814	0.662	0.338
Stimulating your mentees' creativity	0.795	0.632	0.368
Acknowledging your mentees' professional contributions	0.739	0.546	0.454
Negotiating a path to professional independence with your mentees	0.825	0.681	0.319
Taking into account the biases and prejudices you bring to your mentor/mentee relationship	0.789	0.623	0.377
Working effectively with mentees whose personal background is different from your own (age, race, gender, class, region, culture, religion, family composition etc.)	0.723	0.522	0.478
Helping your mentees network effectively	0.694	0.481	0.519
Helping your mentees set career goals	0.831	0.691	0.309
Helping your mentees balance work with their personal life	0.695	0.483	0.517
Understanding your impact as a role model	0.826	0.682	0.318
Helping your mentees acquire resources (e.g. grants, etc.)	0.666	0.444	0.556

Appendix IV
Parallel Analysis Scree Plots



Appendix V
EFA Oblique Rotation Results

	1-factor	4-factor				5-factor					6-factor					
I1	0.707	-0.03	0.833	0.003	0.036	-0.01	-0.03	0.823	0.013	0.079	-0.02	0.814	0.008	-0.03	0.084	0.022
I2	0.761	-0.06	0.645	0.093	0.284	0.014	0.17	0.755	0.09	-0.13	-0.06	0.743	0.107	0.142	-0.1	0.124
I3	0.732	0.134	0.711	-0.01	0.015	0.148	-0.07	0.739	0.012	0.045	0.166	0.76	-0.02	-0	0.008	-0.03
I4	0.768	0.101	0.568	0.251	-0.03	-0.04	0.063	0.408	0.242	0.347	0.066	0.417	0.194	0.111	0.304	-0.15
I5	0.808	0.123	0.525	0.327	-0.04	-0.01	0.04	0.378	0.323	0.314	0.089	0.388	0.28	0.081	0.275	-0.13
I6	0.661	0.108	0.017	0.659	-0.01	0.089	-0	0.008	0.661	0.035	0.186	0.022	0.629	0.074	-0.02	-0.13
I7	0.738	-0	0.007	0.867	0.024	0.026	-0.02	0.023	0.906	-0.06	0.047	0.031	0.89	-0.02	-0.06	-0.01
I8	0.809	0.023	0.062	0.784	0.121	0.009	0.137	0.039	0.774	0.036	-0.04	0.028	0.804	0.064	0.083	0.081
I9	0.765	0.278	0.306	0.278	0.015	0.088	0.143	0.145	0.268	0.373	-0.05	0.12	0.316	0.003	0.491	0.16
I10	0.731	0.198	0.093	0.115	0.543	0.201	0.53	0.127	0.107	-0.06	-0.04	0.112	0.176	0.387	0.046	0.342
I11	0.828	0.348	0.12	0.253	0.271	0.262	0.316	0.08	0.252	0.121	0.115	0.076	0.287	0.235	0.189	0.194
I12	0.658	0.038	0.018	0.104	0.771	-0.01	0.859	-0.02	0.054	0.018	0.002	-0.03	0.005	0.956	-0.02	-0.03
I13	0.685	0.028	0.075	0.072	0.794	-0.01	0.871	0.038	0.023	0.015	-0.02	0.032	-0.02	0.945	-0.01	-0.01
I14	0.847	0.429	0.126	0.122	0.341	0.294	0.432	0.051	0.112	0.197	0.101	0.04	0.16	0.317	0.291	0.251
I15	0.789	0.682	-0.01	0.099	0.115	0.522	0.193	-0.06	0.113	0.213	0.288	-0.07	0.172	0.08	0.326	0.299
I16	0.814	0.655	0.146	-0.02	0.14	0.558	0.15	0.153	0.002	0.111	0.325	0.158	0.053	0.066	0.2	0.3
I17	0.795	0.584	0.101	0.082	0.135	0.526	0.118	0.138	0.109	0.041	0.333	0.146	0.148	0.068	0.107	0.246
I18	0.739	0.559	0.082	-0.09	0.325	0.583	0.251	0.198	-0.07	-0.12	0.226	0.204	0.016	0.087	0.003	0.514
I19	0.825	0.584	-0.03	0.127	0.284	0.582	0.234	0.063	0.154	-0.09	0.407	0.083	0.181	0.226	-0.05	0.228
I20	0.789	0.454	0.37	0.099	-0.05	0.185	0.127	0.162	0.076	0.53	0.141	0.16	0.076	0.075	0.57	0.033
I21	0.723	0.522	0.418	-0.09	-0.06	0.293	0.072	0.259	-0.11	0.453	0.252	0.264	-0.12	0.064	0.469	0.025
I22	0.694	0.778	-0.04	0.115	-0.13	0.717	-0.17	0.017	0.165	0.039	0.727	0.056	0.131	-0.03	-0.03	-0.01
I23	0.831	0.812	0.008	-0.03	0.131	0.744	0.103	0.075	0.011	0.025	0.547	0.098	0.039	0.107	0.069	0.237
I24	0.695	0.727	0.038	0.043	-0.07	0.523	0.037	-0.06	0.054	0.327	0.56	-0.04	0.004	0.146	0.291	-0.08
I25	0.826	0.749	0.088	0.158	-0.12	0.592	-0.06	0.05	0.187	0.23	0.564	0.079	0.162	0.013	0.21	0.009
I26	0.666	0.749	-0.03	-0.01	0.009	0.731	-0.06	0.06	0.036	-0.04	0.661	0.096	0.02	0.044	-0.07	0.085
df	299	227				205					184					
chisq	1678.51	701.95				575.38					451.99					
RMSE	0.052	0.026				0.022					0.019					
TLI	0.849	0.931				0.941					0.952					
RMSEA	0.092	0.058				0.058					0.052					
90% lower	0.086	0.051				0.051					0.045					
90% upper	0.095	0.062				0.062					0.057					

Appendix VI
*Target Rotation Results and Correlations Between Sub-domains**

	Communication	Expectation	Assessing	Independent	Diversity	Profession
I1	0.93	-0.04	-0.08	0.04	0.04	-0.09
I2	0.84	0.06	0.09	0.14	-0.14	-0.15
I3	0.87	-0.12	-0.04	0	-0.02	0.12
I4	0.55	0.14	0.1	-0.16	0.24	0.1
I5	0.52	0.22	0.07	-0.13	0.2	0.12
I6	0.14	0.51	0.08	-0.14	-0.08	0.25
I7	0.15	0.79	-0.02	-0.02	-0.15	0.09
I8	0.13	0.77	0.06	0.07	-0.02	-0.02
I9	0.16	0.41	0	0.2	0.38	-0.08
I10	0.1	0.25	0.37	0.38	-0.02	-0.13
I11	0.1	0.32	0.22	0.24	0.11	0.08
I12	0	-0.01	0.97	-0.07	-0.03	0.02
I13	0.07	-0.03	0.95	-0.05	-0.03	-0.01
I14	0.03	0.24	0.31	0.31	0.21	0.05
I15	-0.11	0.26	0.07	0.4	0.24	0.23
I16	0.13	0.1	0.04	0.41	0.13	0.24
I17	0.14	0.16	0.05	0.34	0.05	0.27
I18	0.13	0.1	0.04	0.63	-0.05	0.08
I19	0.08	0.15	0.21	0.31	-0.09	0.35
I20	0.2	0.14	0.07	0.08	0.48	0.14
I21	0.29	-0.08	0.05	0.09	0.41	0.23
I22	0.08	0.01	-0.04	0.08	-0.04	0.74
I23	0.07	0.02	0.09	0.35	0.03	0.48
I24	-0.02	-0.04	0.16	-0.01	0.26	0.6
I25	0.11	0.1	0.01	0.1	0.16	0.57
I26	0.09	-0.07	0.03	0.19	-0.08	0.64

Corr	comm	exp	asses	indep	div	profes
comm	1					
exp	0.68	1				
asses	0.61	0.6	1			
indep	0.6	0.44	0.59	1		
div	0.47	0.36	0.26	0.32	1	
profes	0.67	0.63	0.48	0.56	0.46	1

*Blue represents higher values, and red lower value

Appendix VII

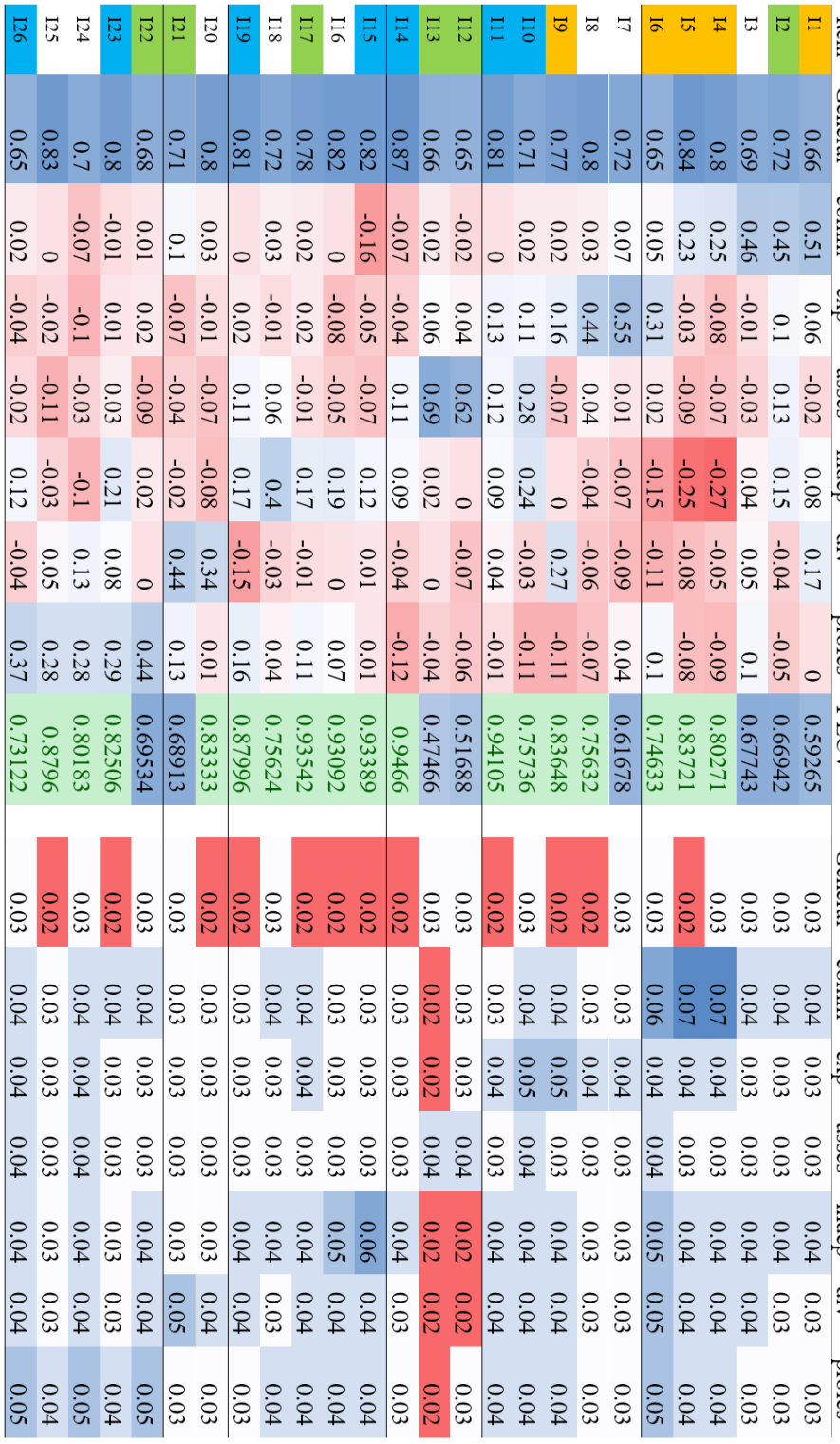
*EFA Bi-factor Models**

	1-f	5 sub-domains						I-ECV	6 sub-domains							I-ECV
I1	0.707	0.713	-0.072	-0.032	0.438	-0.066	-0.049	0.713	0.719	-0.057	-0.006	-0.060	0.453	0.042	0.095	0.699
I2	0.761	0.748	0.114	0.066	0.428	-0.063	0.102	0.722	0.769	0.103	0.030	-0.101	0.379	0.007	-0.112	0.769
I3	0.732	0.727	-0.060	-0.034	0.415	0.072	-0.056	0.741	0.738	-0.057	-0.040	0.061	0.376	0.103	0.005	0.772
I4	0.768	0.775	-0.027	0.043	0.175	-0.040	-0.255	0.857	0.780	-0.032	-0.002	-0.039	0.082	0.419	0.025	0.767
I5	0.808	0.812	-0.037	0.095	0.159	-0.026	-0.228	0.882	0.819	-0.049	0.050	-0.030	0.061	0.396	-0.011	0.801
I6	0.661	0.643	0.046	0.370	0.007	0.098	-0.104	0.722	0.638	0.053	0.361	0.104	-0.013	0.117	-0.028	0.719
I7	0.738	0.726	0.026	0.522	0.016	0.001	-0.006	0.658	0.723	0.039	0.557	0.008	0.021	-0.004	-0.032	0.625
I8	0.809	0.805	0.064	0.432	-0.021	-0.095	0.009	0.764	0.801	0.078	0.433	-0.088	-0.033	0.010	-0.015	0.760
I9	0.765	0.795	-0.095	0.030	-0.068	-0.191	-0.083	0.916	0.774	-0.052	0.109	-0.139	-0.043	-0.013	0.275	0.843
I10	0.731	0.722	0.289	0.048	0.004	-0.104	0.237	0.774	0.731	0.286	0.029	-0.129	-0.018	-0.165	-0.091	0.798
I11	0.828	0.822	0.122	0.080	-0.045	-0.028	0.065	0.960	0.818	0.135	0.092	-0.022	-0.054	-0.060	0.030	0.951
I12	0.658	0.634	0.669	0.028	-0.020	-0.007	-0.002	0.472	0.628	0.648	0.024	-0.017	-0.039	0.016	-0.044	0.481
I13	0.685	0.661	0.660	0.009	0.015	-0.022	0.012	0.500	0.650	0.705	0.023	-0.009	0.024	-0.023	0.016	0.458
I14	0.847	0.850	0.156	-0.025	-0.095	-0.063	0.074	0.943	0.858	0.146	-0.064	-0.093	-0.173	0.019	-0.047	0.917
I15	0.789	0.803	-0.035	-0.052	-0.189	0.036	0.097	0.928	0.806	-0.038	-0.069	0.025	-0.242	-0.050	-0.002	0.906
I16	0.814	0.817	-0.031	-0.093	-0.032	0.086	0.140	0.947	0.827	-0.038	-0.121	0.063	-0.078	-0.072	-0.047	0.953
I17	0.795	0.789	-0.007	-0.006	-0.011	0.116	0.133	0.952	0.796	-0.008	-0.020	0.099	-0.037	-0.093	-0.044	0.966
I18	0.739	0.739	0.055	-0.085	0.021	0.047	0.378	0.778	0.758	0.048	-0.116	0.003	-0.002	-0.282	-0.138	0.834
I19	0.825	0.803	0.134	0.055	-0.016	0.198	0.184	0.872	0.817	0.121	0.005	0.161	-0.064	-0.084	-0.168	0.893
I20	0.789	0.821	-0.109	-0.125	-0.063	-0.074	-0.205	0.895	0.792	-0.057	-0.030	0.004	-0.032	0.075	0.353	0.822
I21	0.723	0.749	-0.118	-0.217	0.012	0.021	-0.172	0.860	0.721	-0.056	-0.117	0.105	0.080	-0.002	0.423	0.709
I22	0.694	0.681	-0.111	0.025	-0.032	0.417	-0.011	0.711	0.677	-0.094	0.042	0.429	-0.022	-0.032	0.025	0.699
I23	0.831	0.823	-0.002	-0.074	-0.048	0.256	0.145	0.878	0.821	0.019	-0.046	0.265	-0.025	-0.187	0.038	0.860
I24	0.695	0.700	-0.049	-0.105	-0.136	0.259	-0.173	0.792	0.678	-0.015	-0.057	0.310	-0.124	0.090	0.185	0.746
I25	0.826	0.826	-0.114	-0.010	-0.062	0.260	-0.085	0.881	0.816	-0.093	0.010	0.285	-0.073	0.059	0.088	0.862
I26	0.666	0.649	-0.037	-0.034	-0.004	0.380	0.085	0.732	0.654	-0.034	-0.040	0.367	-0.010	-0.078	-0.056	0.744
df	299	184							164							
chisq	1678.51	451.99							346.810							
RMSE	0.052	0.019							0.016							
TLI	0.849	0.952							0.963							
RMSEA	0.092	0.052							0.046							
lower	0.086	0.045							0.038							
upper	0.095	0.057							0.051							

*Blue represents higher values, and red lower values

Appendix VIII

*Bi-factor Model with 6 Sub-domain EFA under Target Rotation, and Residuals**



Item	General	comm	exp	asses	indep	div	profes	I-ECV	General	comm	exp	asses	indep	div	profes
I1	0.66	0.51	0.06	-0.02	0.08	0.17	0	0.59265	0.03	0.04	0.03	0.03	0.04	0.03	0.03
I2	0.72	0.45	0.1	0.13	0.15	-0.04	-0.05	0.66942	0.03	0.04	0.03	0.03	0.04	0.03	0.03
I3	0.69	0.46	-0.01	-0.03	0.04	0.05	0.1	0.67743	0.03	0.04	0.03	0.03	0.04	0.04	0.03
I4	0.8	0.25	-0.08	-0.07	-0.27	-0.05	-0.09	0.80271	0.03	0.07	0.04	0.03	0.04	0.04	0.04
I5	0.84	0.23	-0.03	-0.09	-0.25	-0.08	-0.08	0.83721	0.02	0.07	0.04	0.03	0.04	0.04	0.04
I6	0.65	0.05	0.31	0.02	-0.15	-0.11	0.1	0.74633	0.03	0.06	0.04	0.04	0.05	0.05	0.05
I7	0.72	0.07	0.55	0.01	-0.07	-0.09	0.04	0.61678	0.03	0.03	0.04	0.03	0.03	0.03	0.03
I8	0.8	0.03	0.44	0.04	-0.04	-0.06	-0.07	0.75632	0.02	0.03	0.04	0.03	0.03	0.03	0.03
I9	0.77	0.02	0.16	-0.07	0	0.27	-0.11	0.83648	0.02	0.04	0.05	0.03	0.04	0.04	0.04
I10	0.71	0.02	0.11	0.28	0.24	-0.03	-0.11	0.75736	0.03	0.04	0.05	0.04	0.04	0.04	0.04
I11	0.81	0	0.13	0.12	0.09	0.04	-0.01	0.94105	0.02	0.03	0.04	0.03	0.04	0.04	0.04
I12	0.65	-0.02	0.04	0.62	0	-0.07	-0.06	0.51688	0.03	0.03	0.03	0.04	0.02	0.02	0.03
I13	0.66	0.02	0.06	0.69	0.02	0	-0.04	0.47466	0.03	0.02	0.02	0.04	0.02	0.02	0.02
I14	0.87	-0.07	-0.04	0.11	0.09	-0.04	-0.12	0.9466	0.02	0.03	0.03	0.03	0.04	0.03	0.03
I15	0.82	-0.16	-0.05	-0.07	0.12	0.01	0.01	0.93389	0.02	0.03	0.03	0.03	0.06	0.04	0.04
I16	0.82	0	-0.08	-0.05	0.19	0	0.07	0.93092	0.02	0.03	0.03	0.03	0.05	0.04	0.04
I17	0.78	0.02	0.02	-0.01	0.17	-0.01	0.11	0.93542	0.02	0.04	0.04	0.03	0.04	0.04	0.04
I18	0.72	0.03	-0.01	0.06	0.4	-0.03	0.04	0.75624	0.03	0.04	0.03	0.03	0.04	0.03	0.04
I19	0.81	0	0.02	0.11	0.17	-0.15	0.16	0.87996	0.02	0.03	0.03	0.03	0.04	0.04	0.03
I20	0.8	0.03	-0.01	-0.07	-0.08	0.34	0.01	0.83333	0.02	0.03	0.03	0.03	0.03	0.04	0.03
I21	0.71	0.1	-0.07	-0.04	-0.02	0.44	0.13	0.68913	0.03	0.03	0.03	0.03	0.03	0.05	0.03
I22	0.68	0.01	0.02	-0.09	0.02	0	0.44	0.69534	0.03	0.04	0.03	0.03	0.04	0.04	0.05
I23	0.8	-0.01	0.01	0.03	0.21	0.08	0.29	0.82506	0.02	0.04	0.03	0.03	0.03	0.03	0.04
I24	0.7	-0.07	-0.1	-0.03	-0.1	0.13	0.28	0.80183	0.03	0.04	0.04	0.04	0.04	0.04	0.05
I25	0.83	0	-0.02	-0.11	-0.03	0.05	0.28	0.8796	0.02	0.03	0.03	0.03	0.03	0.03	0.04
I26	0.65	0.02	-0.04	-0.02	0.12	-0.04	0.37	0.73122	0.03	0.04	0.04	0.04	0.04	0.04	0.05

*Blue represents higher values, and red lower values

Appendix IX
CFA Results

	1-factor	6-factor						Bi-factor Model with 6 Sub-domain						
I1	0.66	0.70	0	0	0	0	0	0.64	0.44	0	0	0	0	0
I2	0.70	0.75	0	0	0	0	0	0.68	0.37	0	0	0	0	0
I3	0.67	0.71	0	0	0	0	0	0.65	0.44	0	0	0	0	0
I4	0.73	0.78	0	0	0	0	0	0.71	0.29	0	0	0	0	0
I5	0.77	0.82	0	0	0	0	0	0.75	0.28	0	0	0	0	0
I6	0.62	0.66	0	0	0	0	0	0.63	0.01	0	0	0	0	0
I7	0.69	0	0.73	0	0	0	0	0.69	0	0.33	0	0	0	0
I8	0.77	0	0.81	0	0	0	0	0.75	0	0.75	0	0	0	0
I9	0.73	0	0.77	0	0	0	0	0.73	0	0.13	0	0	0	0
I10	0.67	0	0.71	0	0	0	0	0.67	0	0.08	0	0	0	0
I11	0.80	0	0.83	0	0	0	0	0.80	0	0.07	0	0	0	0
I12	0.60	0	0	0.71	0	0	0	0.59	0	0	0.80	0	0	0
I13	0.63	0	0	0.75	0	0	0	0.62	0	0	0.56	0	0	0
I14	0.80	0	0	0.95	0	0	0	0.81	0	0	0.16	0	0	0
I15	0.75	0	0	0	0.78	0	0	0.75	0	0	0	0.14	0	0
I16	0.78	0	0	0	0.82	0	0	0.78	0	0	0	0.31	0	0
I17	0.75	0	0	0	0.79	0	0	0.75	0	0	0	0.24	0	0
I18	0.68	0	0	0	0.71	0	0	0.67	0	0	0	0.29	0	0
I19	0.79	0	0	0	0.82	0	0	0.79	0	0	0	0.18	0	0
I20	0.76	0	0	0	0	0.86	0	0.76	0	0	0	0	0.44	0
I21	0.68	0	0	0	0	0.77	0	0.67	0	0	0	0	0.35	0
I22	0.66	0	0	0	0	0	0.72	0.64	0	0	0	0	0	0.42
I23	0.79	0	0	0	0	0	0.86	0.78	0	0	0	0	0	0.28
I24	0.66	0	0	0	0	0	0.72	0.64	0	0	0	0	0	0.33
I25	0.79	0	0	0	0	0	0.86	0.77	0	0	0	0	0	0.29
I26	0.63	0	0	0	0	0	0.69	0.61	0	0	0	0	0	0.40
General								1	0	0	0	0	0	0
f1		1	0.90	0.76	0.84	0.83	0.82	0	1	0	0	0	0	0
f2		0.90	1	0.86	0.89	0.80	0.81	0	0	1	0	0	0	0
f3		0.76	0.86	1	0.83	0.68	0.70	0	0	0	1	0	0	0
f4		0.84	0.89	0.83	1	0.81	0.91	0	0	0	0	1	0	0
f5		0.83	0.80	0.68	0.81	1	0.84	0	0	0	0	0	1	0
f6		0.82	0.81	0.70	0.91	0.84	1	0	0	0	0	0	0	1

Appendix X
Eight-item Short Form

Communication	Establishing a relationship based on trust
Expectation	Aligning your expectations with your mentees'
Assessing	Accurately estimating your mentees' level of scientific knowledge
Independence	Building mentees' confidence
Independence	Stimulating your mentees' creativity
Diversity	Taking into account the biases and prejudices you bring to your men-tor/mentee relationship
Profession	Helping your mentees balance work with their personal life
Profession	Understanding your impact as a role model

Appendix 11

If the simplified version of previous procedure is repeated without considering the sub-constructs, based on different cut-offs, the short forms (8-item version based on 1-factor solution and 12-item version based on I-ECV) are presented below.

	Item	14-Items (alpha=.94)	8 Items (alpha=.93)	12 Items (alpha=.95)
Communication	Active listening			
	Providing constructive feedback	.94		
	Establishing a relationship based on trust	.94		
	Identifying and accommodating different communication styles			
	Employing strategies to improve communication with mentees		.93	.95
	Coordinating effectively with your mentees' other mentor			
Expectation	Working with mentees to set clear expectations of the men-toring relationship	.94	.93	
	Aligning your expectations with your mentees'	.93		
	Considering how personal and professional differences may impact expectations			.95
	Working with mentees to set research goals			
	Helping mentees to develop strategies to meet goals		.92	.95
Assessing	Accurately estimating your mentees' level of scientific knowledge	.94		
	Accurately estimating your mentees' ability to conduct research	.94		
	Employing strategies to enhance your mentees' knowledge and abilities		.92	.95
Independence	Motivating your mentees			.95
	Building mentees' confidence	.93	.92	.95
	Stimulating your mentees' creativity	.93		.95
	Acknowledging your mentees' professional contributions	.94		.95
	Negotiating a path to professional independence with your mentees		.92	.95
Diversity	Taking into account the biases and prejudices you bring to your mentor/mentee relationship	.94		.95
	Working effectively with mentees whose personal back-ground is different from your own (age, race, gender, class, region, culture, religion, family composition etc.)	.94		
Profession	Helping your mentees network effectively	.94		
	Helping your mentees set career goals		.92	.95
	Helping your mentees balance work with their personal life	.94		
	Understanding your impact as a role model	.93	.93	.95
	Helping your mentees acquire resources (e.g. grants, etc.)			

Appendix 12
TIC Comparisons