On the Psychometric Properties of the Multigroup Ethnic Identity Measure—Revised (MEIM-R)

Anthony D. Ong[1]

Cornell University

Jean S. Phinney

California State University, Los Angeles

Letters to the Editor, *Human Development*, September 10, 2014 [vol. 57 (4)]

[1] To whom correspondence should be addressed. Email: ado4@cornell.edu. Department of Human Development, G77 Martha Van Rensselaer Hall, Cornell University, Ithaca, NY 14853-4401, Tel: (607) 255-9993; Fax: (607) 255-9856

On the Psychometric Properties of the Multigroup Ethnic Identity Measure—Revised (MEIM-R)

In his conceptual and methodological critique of ethnicity research in developmental psychology, Gjerde (2014) questioned the measurement validity of the Multigroup Ethnic Identity Measure—Revised (MEIM-R) and maintained that Phinney and Ong (2007) "try to promote a one-factor model as a two-factor model" (Gjerde, 2014, p. 181). The research findings said to support this conclusion, however, are inadequate. The summary of evidence is incomplete and the claim that the MEIM-R is "not psychometrically sound" omits much of the extant evidence. In this commentary, we emphasize the need for a more balanced consideration of the meaning of ethnic identity and suggest that the evidentiary foundation on which to draw conclusions regarding the factor structure of the six-item MEIM-R, in particular, should rest on replication studies across multiple datasets. We elaborate on these points below.

To date, six factor-analytic studies of the six-item MEIM-R, including Phinney and Ong (2007), have been published. The analytic sample sizes of these studies range from 106 to 4,311 and comprise a broad spectrum of children, adolescents, and adults (age range, 12-59 yrs) from diverse ethnic and racial backgrounds (see Table 1 for study characteristics). In what follows, we summarize the results from validation studies, published since Phinney and Ong (2007), that evaluate the psychometric properties of the MEIM-R, including its factor structure and measurement invariance. Importantly, in each of the studies reviewed, the fit of competing models (e.g., one-factor structure, correlated-two factor structure) were evaluated using confirmatory factor analysis (CFA). To evaluate model fit, we include multiple fit indices: The comparative fit index (CFI), the standardized root-mean-square residual (SRMR), the root-mean-square error of approximation (RMSEA), and the 90% confidence interval (CI) for RMSEA. In general, a good model fit is characterized by CFI values close to or greater than .95, SRMR

values close to or less than .08, and RMSEA values close to or less than .06. An acceptable model fit includes CFI values close to or greater than .90 and RMSEA and SRMR values close to or less than .10 (Browne & Cudeck, 1992; Hu & Bentler, 1999).

Factor Structure

Yoon (2011) conducted separated CFAs on a sample of European American (*n* = 100; age range, 19-59 yrs) and minority (*n* = 100; age range, 20-64 yrs) counseling students. Specifically, the fit of three competing models was evaluated: a one-factor model in which all six MEIM-R items loaded onto a single latent factor; an uncorrelated two-factor model in which latent constructs of *exploration* and *commitment* were unrelated to each other; and the theoretical two-factor model in which exploration and commitment factors were allowed to correlate. In comparison to the one-factor model and uncorrelated two-factor model, the correlated two-factor model showed the best fit, leading Yoon (2011) to conclude "the correlated two-factor structure of the MEIM-R (exploration, commitment) was confirmed in both groups" (p. 150).

Ashdown, Homa, and Brown (2014) also evaluated the fit of the one-factor, uncorrelated two-factor, and correlated two-factor structure in a sample of adolescents (*N* = 247; age range, 16-18 yrs). Similar to Yoon (2011), the theoretical, correlated two-factor model representing exploration and commitment fit the data better than the alternative one-factor or uncorrelated two-factor models, respectively. Additionally, the authors tested whether an adapted version of the MEIM-R could measure exploration and commitment in adolescents' gender and religious identities. Here too, CFAs for the correlated two-factor models representing gender and religious identity demonstrated a better fit over the alternative models.

Brown et al. (2014) examined the dimensionality of ethnic identity in a large sample (*N* = 1,463; age range, 18-50 yrs) of racial and ethnic minorities, including Asian (*n* = 630),

Black/African American (*n* = 58), Hispanic (*n* = 240), multiethnic (*n* = 160), and White participants (*n* = 375). On the basis of CFA, Brown et al. (2014) concluded, "A one-factor model showed adequate fit according to the SRMR but poor fit according to the RMSEA and CFI. An uncorrelated two-factor model showed overly poor fit. A correlated two-factor model fit best in the entire sample, showing good fit according to the SRMR and CFI and adequate fit according to the RMSEA, and the best fit, i.e., lowest AIC value relative to other models" (p. 157).

Finally, two recent studies—one based on a sample of East Asian adolescents living in Canada (Homma, Zumbo, Saewyc, & Wong, 2014) and another based a clinical sample of Latino adolescents (Burrow-Sanchez, 2014)—also provide relevant evidence on the factor structure of the MEIM-R. Homma et al. (2014) conducted a series of CFAs on a sample of East Asian adolescents from western Canada (*N* = 4,311; age range, 12-19 yrs). The authors noted that although the one-factor model showed an adequate fit to the data according to CFI (.98) and SRMR (.04), it showed a poor fit according to RMSEA (.09). In comparison, the two-factor correlated model demonstrated good model fit (SRMR = .03, RMSEA = .06, CFI = .99). Using the MEIM-R, Burrow-Sanchez (2014) evaluated the fit of three competing factor structure models (i.e., a 12-item one-factor model, a 12-item correlated two-factor model, and a six-item correlated two-factor model) in a clinical sample of Latino adolescents (*N* = 106; age range, 13-18 yrs) with substance use disorders. CFAs indicated that the theoretical, six-item correlated two-factor model provided the best fit to the data.[1]

Measurement Invariance

Evidence that the factor structure of the MEIM-R is equivalent across different subgroups is established through formal tests of measurement invariance. As discussed by Ong, Fuller-

[1] It should be noted that the items in this study were based on Spanish translation and back-translation that resulted in slight modifications in wording from the original MEIM-R.

Rowell, and Phinney (2010), evidence of measurement invariance can be demonstrated by testing a sequence of invariance hypotheses focusing on the items (configural invariance), loadings (metric invariance), intercepts (scalar invariance), and specific-factors (strict invariance). Establishing measurement invariance, thus, provides evidence not only that respondents from different groups can be legitimately compared on the same instrument, but also that observed group mean differences in raw scores reflect valid and meaningful group differences at the level of the latent variable(s) assumed to underlie those scores (Meredith, 1993). Of the five studies reviewed, three examined the invariance of the theoretical correlated two-factor model (Brown, et al., 2014; Homma, et al., 2014; Yoon, 2011). All used multigroup confirmatory factor analytic procedures.

Yoon (2011) conducted invariance tests of factor loadings and factor correlations. In comparison to a model with no constraints, a model in which all factor loadings and factor correlations were constrained to be equivalent across samples resulted in negligible decrease in model fit (ΔCFI < .02). Brown et al. (2014) assessed invariance of factor structure, factor loadings, and item intercepts across five racial/ethnic groups and found provisional support for measurement invariance of the correlated two-factor model. Finally, Homma et al. (2014) evaluated strict measurement invariance by constraining the factor structure, factor loadings, intercepts, and residual variances of the MEIM-R to be equal across age groups. Support for strict invariance was found across early, middle, and late adolescence, leading the authors to conclude that the “items were interpreted similarly across age groups” (p. 10).

To summarize, with respect to measures of ethnic identity, the principal thesis advanced by Gjerde (2014) is that the MEIM-R is psychometrically unsound. Analysis of the available evidence, however, indicates that this thesis lacks empirical support. Indeed, contrary to claim,

the MEIM-R demonstrates acceptable to good psychometric properties across diverse and independent samples. This is indicated by factor analytic evidence (i.e., factor structure, measurement invariance) from the five studies, as summarized in this commentary. More fundamentally, whether or not Gjerde's thesis *is* correct, the argument for it in his critique does not show it to be correct. That argument lacks the evidence required to uphold the thesis. This should be seen by those who do not like the thesis, as well as by those who would champion it. Gjerde's (2014) other criticisms of the MEIM, namely that it (a) "fails to address a full perspective on ethnicity, lacking consideration such as power, inequality, class, religion, gender, sexual orientation, age, and generation" (p. 181) and (b) "cannot gauge ethnic identities as they emerge in the minutiae of everyday encounters" (p. 181) are clearly unwarranted, as illustrated by research cited in this commentary (e.g., Ashdown, et al., 2014) as well as those cited in the commentary by Kiang (2014, this issue). Finally, Gjerde's (2014) portrayal—that "MEIM studies only rely on self-report data and are therefore susceptible to response biases such as impression management, self-deception, and acquiescent responding" (p. 181)—is a caricature. It is a critique that is neither unique to studies that use the MEIM or to those using measures that Gjerde ostensibly privileges over the MEIM-R, e.g., Sellers' (1997) Multidimensional Inventory of Black Identity (MIBI) and Umaña-Taylor's (2004) Ethnic Identity Scale (EIS).

Concluding Remarks

In closing, we underscore that further validation of the MEIM-R represents a main priority for future research. Adjudicating the measurement validity of the MEIM-R and competing measures will require piecing together multiple sources of evidence (e.g., factorial validity, sociodemographic and psychosocial correlates, underlying mechanisms), each providing strong empirical ties to theory-driven constructs. From a developmental perspective, there is also

a pressing need for studies that examine the dynamic nature of ethnic identity using approaches that allow for a mapping of “the person” as a highly variable unit (Ong, et al., 2010). Here, we are in agreement with Gjerde (2014): “To understand this variability is one of the most important tasks awaiting DP [developmental psychology] ethnicity researchers” (p. 200).

Acknowledgments

We thank Nancy Wells for helpful feedback on a previous version of this commentary. We also extend thanks to Brien Ashdown, Carrie Brown, Jason Burrow-Sanchez, and Natalie Homa for making available unpublished model fit statistics reported in Table 1.

Table 1
Fit Indices of Confirmatory Factor Analyses for the Six-Item MEIM-R

Author(s)	Sample	Age	Model	CFI	SRMR	RMSEA [90% CI]
Ong & Phinney (2007)*	241	17-25	One factor	0.91	0.16	0.09 [.08-.12]
			Uncorrelated 2 factors	0.87	0.36	0.23 [.20-.25]
			Correlated 2 factors	0.98	0.05	0.04 [.03-.06]
Yoon (2011)	100 E	19-59	One factor	0.86	0.11	0.28 [.22-.34]
			Uncorrelated 2 factors	0.89	0.33	0.22 [.17-.28]
			Correlated 2 factors	0.97	0.04	0.12 [.05-.19]
	189 M	20-34	One factor	0.80	0.12	0.38 [.34-.42]
			Uncorrelated 2 factors	0.88	0.30	0.22 [.18-.26]
			Correlated 2 factors	0.98	0.05	0.12 [.07-.17]
Ashdown et al. (2014)	247	16-18	One factor	0.90	0.06	0.14 [.11-.18]
			Uncorrelated 2 factors	0.75	0.26	0.23 [.19-.26]
			Correlated 2 factors	0.97	0.04	0.09 [.05-.13]
Brown et al. (2014)*	1,463	18-50	One factor	0.82	0.08	0.20 [.18-.21]
			Uncorrelated 2 factors	0.76	0.29	0.22 [.21-.24]
			Correlated 2 factors	0.98	0.03	0.07 [.06-.09]
Homma et al. (2014)*	4,311	12-19	One factor	0.98	0.04	0.09 [.08-.10]
			Correlated 2 factors	0.99	0.03	0.06 [.05-.07]
Burrow-Sanchez (2014)	106	13-18	One factor†	0.78	0.09	0.15 [.13-.17]
			Correlated 2 factors†	0.84	0.08	0.13 [.10-.18]
			Correlated 2 factors	0.94	0.05	0.11 [.04-.17]

Note. E = European American; M=Minority; *Used Satorra-Benter chi-square; CFI=comparative fit index; SRMR=Standardized root-mean-square residual; RMSEA=root-mean-square error of approximation; CI=confidence interval. †Used 12-item MEIM.

References

Ashdown, B. K., Homa, N., & Brown, C. M. (2014). Measuring gender identity and religious identity with adapted versions of the Multigroup Ethnic Identity Measure—Revised. *Journal of Educational and Developmental Psychology, 4*, 226-237.

Brown, S. D., Unger, K. A., Mevi, A. A., Hedderson, M. M., Shan, J., Quesenberry, C. P., & Ferrara, A. (2014). The multigroup ethnic identity measure-revised: Measurement invariance across racial and ethnic groups. *Journal of Counseling Psychology, 61*, 154-161.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods and Research, 21*, 230–258.

Burrow-Sanchez, J. J. (2014). Measuring ethnic identity in Latino adolescents with substance use disorders. *Substance Use & Misuse, 49*, 982-986.

Gjerde, P. F. (2014). An evaluation of ethincity research in developmental psychology: Critiques and recommendations. *Human Development, 57*, 176-205.

Homma, Y., Zumbo, B. D., Saewyc, E. M., & Wong, S. T. (2014). Psychometric evaluation of the six-item version of the Multigroup Ethnic Identity Measure with East Asian adolescents in Canada. *Identity: An International Journal of Theory and Research, 14*, 1-18.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Kiang, L. (2014). Ethnicity and ethnic identity in context. *Human Development, 57*, 213-221.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.

Ong, A. D., Fuller-Rowell, T. E., & Phinney, J. S. (2010). Measurement of ethnic identity: Recurrent and emergent issues. *Identity: An International Journal of Theory and Research, 10*, 39-49. doi:10.1080/15283481003676226

Phinney, J. S., & Ong, A. D. (2007). Conceptualization and measurement of ethnic identity: Current status and future directions. *Journal of Counseling Psychology, 54*, 271-281. doi:10.1037/0022-0167.54.3.271

Sellers, R. M., Rowley, S. A. J., Chavous, T. M., Shelton, J. N., & Smith, M. A. (1997). Multidimensional Inventory of Black Identity: A preliminary investigation of reliability and constuct validity. *Journal of Personality and Social Psychology, 73*, 805-815.

Umaña-Taylor, A. J., Yazedjian, A., & Bamaca Gomez, M. (2004). Developing the ethnic identity scale using Eriksonian and social identity perspectives. *Identity: An International Journal of Theory and Research, 4*, 9-38.

Yoon, E. (2011). Measuring ethnic identity in the Ethnic Identity Scale and the Multigroup Ethnic Identity Measure-Revised. *Cultural Diversity and Ethnic Minority Psychology, 17*, 144-155.