

Reply

Reply to Hammer, M.R. A Response to Puntì and Dingel's Critique of the Validity of the Intercultural Development Inventory for BIPOC Students. Comment on "Puntì, G.; Dingel, M. Rethinking Race, Ethnicity, and the Assessment of Intercultural Competence in Higher Education. *Educ. Sci.* 2021, 11, 110"

Gemma Puntì ^{1,*} and Molly Dingel ^{2,*} 

¹ College of Individualized and Interdisciplinary Studies, Metropolitan State University, Saint Paul, MN 55106, USA

² Center for Learning Innovation, University of Minnesota Rochester, Rochester, MN 55904, USA

* Correspondence: gemma.puntì@metrostate.edu (G.P.); dinge016@r.umn.edu (M.D.)



Citation: Puntì, G.; Dingel, M. Reply to Hammer, M.R. A Response to Puntì and Dingel's Critique of the Validity of the Intercultural Development Inventory for BIPOC Students. Comment on "Puntì, G.; Dingel, M. Rethinking Race, Ethnicity, and the Assessment of Intercultural Competence in Higher Education. *Educ. Sci.* 2021, 11, 110". *Educ. Sci.* 2022, 12, 180. <https://doi.org/10.3390/educsci12030180>

Academic Editor: Tabitha Grier-Reed

Received: 26 January 2022

Accepted: 26 February 2022

Published: 4 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Hammer claims that the IDI has been validated for BIPOC in the US, but after careful re-examination of the data we presented, we reiterate that there is no evidence of this validation. The studies provided by Hammer reveal how the development of the IDI instrument and its validation focused on *international* cultural experiences not on cultural diversity *within* the US. There was no validity testing done on a racially diverse sample of U.S. Americans, and our qualitative data questions the validity for BIPOC individuals from the US.

Keywords: BIPOC; validity; IDI; intercultural competence; structural inequality; higher education

Hammer [1] raises several questions regarding our paper [2], including challenging our claims and questioning our findings. We stand behind our published work and believe that his claims are unsupported and misunderstand the nature of our work.

1. Validation of IDI for BIPOC in the US

Hammer claims that the IDI has been validated for Black, Indigenous, and People of Color (BIPOC) in the US, but after careful examination of the data presented, we reiterate that there is no evidence of this validation. Hammer cites just three peer-reviewed articles [3–5] and two sources that have not been peer-reviewed and were completed by ACS Ventures under contract by IDI LLC [6,7].

For the three peer-reviewed articles, we see that diversity within countries was not meaningfully assessed. Hammer et al. [4] demonstrate several phases of testing. Phase 1, which focused on the development of the IDI instrument, used interview data of 40 international participants from 18 countries. Individuals from the US were European American ($n = 12$) and South Asian American ($n = 3$). The rest of the participants were not from the US. Second, sample testing was carried out on 226 individuals; 177 were from the US, with 8 identifying as African American, 2 as Asian American, and 4 as Hispanic American. Even if these individuals are combined into one group, 14 individuals do not allow for statistically sound analysis, and there was no indication or attempt to control for participants' ethnic or racial background. For phase 2, which developed the final 50-item IDI, the authors used a sample of 591 respondents. While some information on the cultural identity of the US participants was asked, the authors stated that the "Ethnic background of respondents from the United States was difficult to assess" (p. 432). Self-identification was used, with 178 US white Americans, 32 US African American, 6 US Latino, and 8 US Asian. Thus, there were a total of 46 self-identified BIPOC individuals, which, once again,

does not yield a statistically large sample, and each individual group is not large enough to draw statistical conclusions. Further and importantly, the authors *did not do* additional testing for race, indicating that “additional tests were run on the IDI, examining the effects of gender, age, education, and social desirability”, *but not on race* [4]. Therefore, validity was not explored for BIPOC in the US.

Paige et al. [5] controlled for their 353 subjects only by enrollment in language classes. *There was no information about racial groups.* Finally, Hammer et al. [3] used an international sample, with information on respondents’ national origin, but those from the US were labeled as either being from the US or as “U.S. American” [3]. *Race was not a variable of analysis.*

Hammer also cites two sources that are not peer reviewed, completed under contract with IDI LLC [6,7]. In our paper, we relied only on peer-reviewed articles because they provide a standard of quality control. Wiley [6] is a review of other data regarding the IDI that illustrates how the validity of the IDI has been questioned by other authors. Matsumoto and Hwang [8] (cited in [4]) point out that the IDI has not been tested using cross-cultural samples within the US, and state how this is a critical limitation of the IDI. Furthermore, Matsumoto and Hwang [8] argued that the “assessments of the factor structure of the IDI have produced inconsistent results in limited samples that do not correspond to the model intended to be measured” [3,5,9], thus questioning the validity of the instrument itself.

Wiley [7] provides additional testing; however, there were weaknesses in this source. First, while “Ethnic minority status” was one of the variables, we are given no additional information about this category, including how many people were in this sample or what country or background the “ethnic minorities” were from. There is simply not enough information in it to draw conclusions about the IDI’s validity among BIPOC from the US.

Finally Hammer [1] provides numbers of a sample size apparently examined, but he does not cite where he obtained those numbers, and we could not find them. Hammer indicates that the IDI was tested across “218,111 international and domestically diverse respondents along with additional testing undertaken with 20,015 respondents who are self-reported they were members of an ethnic minority” (p. 1), but, though they refer to them on three different pages (pp. 1, 2, 5), no citation is provided for these numbers, and they do not appear in the largest study cited, Wiley [7], or the other sources they cited.

All these studies reveal how the development of the IDI instrument and its validation focused on *international* cultural experiences not on cultural diversity *within* the US. In the development of the IDI instrument, the authors selected individuals from several countries [3,4]. Such a strategy does create a widely diverse sample but also contains two important assumptions: first, that one country equals one culture, and second, that culture equals national origin. These studies therefore implicitly assess intercultural sensitivity towards people from other countries not from within one nation. For instance, when Hammer et al. [4] use the phrase “another culture” they mean another country. The paper reads: “Thirty-two percent ($n = 192$) of the respondents never lived *in another culture*, 14% ($n = 81$) lived overseas less than 3 months, 10% ($n = 57$) lived *in another culture* 3–6 months” (p. 432, emphasis added).

In other words, the IDI focuses on international cultural encounters *not* intercultural experiences within the United States or one nation. While stating that this approach is also valid for BIPOC respondents, the author conflates international diversity with diversity in the US. If international diversity is the same as diversity within the US, then the author is implying that a black person who has lived in the US for their whole life has the same intercultural experiences as a black person who grew up in another country AND that black and white people in the US have the same intercultural experiences. However, these assumptions are undermined by data showing deep divisions between how black and white people in the US experience the same settings [10–12] and that there are complicated divisions between black people who have lived in the US for generations and more recent black immigrants to the US [13,14].

2. Validity, Cultural Difference, and the Findings of the Original Study

2.1. Validity

Samuel Messick [15], in a seminal article, defines validity as an “overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores” (p. 741). He then goes on to argue that “validity is not a property of the test or assessment as such, but rather the meaning of the test scores,” which are a function of both the items but also the persons responding and the context of the assessment [15]. Andrew Wiley, who has conducted significant IDI validation, apparently agrees with this definition, indicating that validity is not a dichotomous concept but instead “a matter of degree, and as evidence supporting an intended interpretation or use of a test score accumulates, the degree of confidence in the validity of a give [sic] test score use can increase along with it” [6] (p. 4). Further, echoing statements made by Messick, Wiley also emphasizes, “. . . validity is not something that belongs to a given test. Instead, validity must be considered within the given context and intended use of a test score; evidence must be accumulated for each use of a test and must help build an argument that the given use is appropriate. . . . as different uses for a test are introduced, it would be appropriate to consider what new evidence would be appropriate” [6] (p. 4). In other words, it is inaccurate to make a blanket statement that the IDI is a valid survey. Instead, the evidence on validity suggests that the IDI can be used to facilitate intercultural understanding in some specific contexts for some populations. The IDI was developed and validated with an international context, not with a national focus and not with the intent of assessing the intercultural competence of BIPOC individuals in the US.

Along these lines, as we have demonstrated above, there was no validity testing done on a racially diverse sample of US Americans. Our qualitative data questioned the validity for BIPOC individuals from the US. Given the nature of validity, and the lack of validity testing on this population, interview data are an appropriate source of information to investigate this question [15].

2.2. Methodological Choices

Our research asks whether there is evidence that the IDI is valid for BIPOC individuals in the US. To explore this question, we used an existing data set of interviews that asked students about their cultural challenges and perceptions of diversity on campus. Our research was therefore NOT an investigation into the cultural competence of the students we interviewed but was a query as to whether there was consistency between how these students discussed diversity on campus, their own cultural challenges, and their IDI scores. We therefore collected all examples of students talking about cultural or diversity issues.

2.2.1. Timing of Interviews

While the interviews were not conducted at the exact same time as the administration of the IDI, the IDI survey was carried out at the end of the first semester and most (if not all) the experiences stated in the students’ quotations used for the paper referred to events that happened prior to taking the IDI survey (a pre-semester orientation, arrival on campus, getting to know their roommates and classmates, or an experience in middle school).

2.2.2. Sample

In their article, Hammer questions our sample, and how many students’ interviews comprised our analysis. Our study was a qualitative case study that focused on ALL the students who we interviewed and who took the IDI ($n = 34$). From this sample, we focused specifically on students who the IDI classified as being in denial or polarization ($n = 11$). We analyzed both white and BIPOC students’ descriptions of cultural experiences and compared students who the IDI classified as being in the same stage. Like many qualitative studies, this was a small sample and not generalizable but provided in-depth understanding of the population interviewed. Given the lack of validity testing for this

population, our data questions the validity of the IDI for these students by illustrating the disparate intercultural experiences that BIPOC and white students had while falling in the same IDI stage. We agree, however, that more work must be carried out in this area and, here and in our previous work, call for more research on this population.

2.2.3. Coding

Hammer questions why we did not report inter-rater reliability (IRR). There is not agreement among scholars that quantitative measures of inter-rater reliability are necessary or, in some cases, even desirable for qualitative research [16–18]; indeed in some cases, desire for IRR may be a case of “evaluating qualitative research reports using the standards for quantitative research” [17] (p. 217). One such situation that does not require quantitative measures is when the data are straightforward [18]; simpler coding schemes are inherently more reliable than complex coding schemes [19]. Such was the case with our dataset, where identifying students’ discussion of culture and diversity was relatively straightforward. “Disagreements” in our case amounted to a missed code, instead of outright disagreement about whether a statement was in fact about diversity or a cultural challenge or not. After this first level of coding was completed, Puntí led in analyzing the texts from these relevant codes to better understand how each of these students discussed their experiences of culture and diversity. Analysis was performed with two aims: to understand these students’ description of their experiences and to compare these descriptions with the IDI categories. This was an iterative process, with Dingel subsequently reviewing the analysis, raising discussion and questions when relevant. This more fluid later process is well within established qualitative methods and does not benefit from quantitative measures of IRR [19–21].

2.2.4. Dunning–Kruger Effect

Hammer asks how we controlled for the Dunning–Kruger effect, defined as being where individuals’ lack of expertise is hidden from them; individuals are unaware of “how deficient their expertise is” and they are not able to recognize their own mistakes because of this deficiency [22] (p. 217). There was no interview question that asked the students to reflect on their IDI scores or on their intercultural competence. In other words, the Dunning–Kruger effect does not apply here because these students were not asked to evaluate their own expertise. The interview was extensive and focused on their holistic educational experiences on campus, from academic standing to their experience with university resources to their experience with diversity. Therefore, as the quotations in the study indicate, the students brought up *their own experiences and perspectives* on diversity and cultural challenges. However, throughout history, many scientists have rejected that BIPOC people have insight into their own experiences on race or that they somehow do not understand these experiences; such a rejection has long been decried by racial scholars as part of white supremacy that upholds a racist structure [23–25]. Further, while it is true that faculty interviewing students can affect their responses, there is no evidence that there was an “expectation” or a “common understanding of what faculty wanted to hear”. As the quotations indicate, both BIPOC and white the students revealed their cultural *experiences* on campus. Some indicated not having any experience with diversity, or that “the language barrier is tough. Different foods are—I’m not gonna lie, sometimes they smell bad to us” (Ethan, [2]). Those comments do not seem an attempt to say “what we wanted to hear”.

2.3. Cultural Difference

Our study revealed incongruence between Black and Latino students categorized by the IDI as being in “Denial” and “Reversal” stages and how they talked about their experiences navigating multiple cultures. Despite validity testing *with other populations in other contexts*, our research questions whether the IDI is valid for *this population in this context—BIPOC individuals living as minoritized individuals within a dominantly white culture*.

There is an enormous body of evidence that white and BIPOC individuals experience the cultural world differently in the US [10–12]. Further, a black person who is born in the US and learns to navigate white culture very early in life should not be equated with a white adult traveling abroad, and, for the first time, being exposed to a different culture from their own. It is possible for white people in the US to ignore race; it is not possible for BIPOC individuals to do so [11,26–28]. Our research adds some examples that we believe warrant further investigation with larger samples and tempered messaging about the IDI's validity.

The Developmental Model of Intercultural Sensitivity (DMIS) theory “postulates that individuals with more complex cultural categories are better able to navigate through cultural differences” [4] (p. 217). This means that the higher the IDI score or the stage one is categorized into, the more she/he will be able to navigate cultural differences and, on the contrary, that the lower the IDI stage, the more difficulty she/he will have navigating cultural differences. As much research has shown, BIPOC individuals in the US are obliged to deal with cultural differences daily in a way white individuals are not [27–29]. Does the DMIS take these racial and cultural differences into account? Our study suggests more research is needed on this question.

3. Final Thoughts

We agree with Hammer that addressing social inequality and increasing intercultural competence *and knowledge of structural inequality* is a critical challenge. However, we need strategies and tools that are solidly based in evidence. In our findings are opportunities for improvement in this important area.

Further, given the nature of validity—that it is not something that belongs to a given test but instead must always be considered in context—we also question whether the context of the testing has fundamentally changed. The DMIS was developed during the 1980s and 1990s, and the IDI peer-reviewed testing was published in 2003 and 2011—nineteen and eleven years ago. Our understanding of intercultural understanding has changed significantly in the past 30 years; the political context and national dialogue around race, culture, and inequality has shifted in explicit and substantial ways in the last 20 years. We encourage additional exploration and questioning around what tools and strategies show the best evidence to create awareness of difference within a racial hierarchy and build bridges for a better shared understanding.

Author Contributions: Conceptualization: G.P. and M.D.; writing–review and editing, G.P. and M.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hammer, M.R. A Response to Puntí and Dingel's Critique of the Validity of the Intercultural Development Inventory for BIPOC Students. Comment on Puntí, G.; Dingel, M. Rethinking Race, Ethnicity, and the Assessment of Intercultural Competence in Higher Education. *Educ. Sci.* 2021, 11, 110. *Educ. Sci.* 2022, 12, 176. [\[CrossRef\]](#)
2. Puntí, G.; Dingel, M. Rethinking Race, Ethnicity, and the Assessment of Intercultural Competence in Higher Education. *Educ. Sci.* 2021, 11, 110. [\[CrossRef\]](#)
3. Hammer, M.R. Additional cross-cultural validity testing of the Intercultural Development Inventory. *Int. J. Intercult. Relat.* 2011, 35, 474–487. [\[CrossRef\]](#)
4. Hammer, M.R.; Bennett, M.J.; Wiseman, R. Measuring intercultural sensitivity: The intercultural development inventory. *Int. J. Intercult. Relat.* 2003, 27, 421–443. [\[CrossRef\]](#)

5. Paige, R.; Jacobs-Cassuto, M.; Yershova, Y.; DeJaeghere, J. Assessing intercultural sensitivity: An empirical analysis of the Hammer and Bennett Intercultural Development Inventory. *Int. J. Intercult. Relat.* **2003**, *27*, 467–486. [[CrossRef](#)]
6. Wiley, A. *Intercultural Development Inventory (IDI): Independent Review*; ACS Ventures: Las Vegas, NV, USA, 2016.
7. Wiley, A. *Validation Analysis of the Intercultural Development Inventory (IDI)*; ACS Ventures: Las Vegas, NV, USA, 2017.
8. Matsumoto, D.; Hwang, H.C. Assessing cross-cultural competence: A review of available tests. *J. Cross-Cult. Psychol.* **2013**, *44*, 849–873. [[CrossRef](#)]
9. Greenholtz, J. Does intercultural sensitivity cross cultures? Validity issues in porting instruments across languages and cultures. *Int. J. Intercult. Relat.* **2005**, *29*, 73–89. [[CrossRef](#)]
10. Desilver, D.; Lipka, M.; Fahmy, D. 10 Things We Know about Race and Policing in the U.S. Available online: <https://www.pewresearch.org/fact-tank/2020/06/03/10-things-we-know-about-race-and-policing-in-the-u-s/> (accessed on 24 January 2022).
11. Harwood, S.A.; Mendenhall, R.; Lee, S.S.; Riopelle, C.; Huntt, M.B. Everyday Racism in Integrated Spaces: Mapping the Experiences of Students of Color at a Diversifying Predominantly White Institution. *Ann. Am. Assoc. Geogr.* **2018**, *108*, 1245–1259. [[CrossRef](#)]
12. Lo, C.C.; McCallum, D.M.; Hughes, M.; Smith, G.P.A.; McKnight, U. Racial Differences in College Students' Assessments of Campus Race Relations. *J. Coll. Stud. Dev.* **2017**, *58*, 247–263. [[CrossRef](#)]
13. Alex-Assensoh, Y.M. African Immigrants and African Americans: An Analysis of Voluntary African Immigration and the Evolution of Black Ethnic Politics in America. *Afr. Asian Stud.* **2009**, *8*, 89–124. [[CrossRef](#)]
14. Waters, M.C.; Kasinitz, P.; Asad, A.L. Immigrants and African Americans. *Annu. Rev. Sociol.* **2014**, *40*, 369–390. [[CrossRef](#)]
15. Messick, S. Validity of Psychological Assessment. *Am. Psychol.* **1995**, *50*, 741–749. [[CrossRef](#)]
16. Armstrong, D.; Gosling, A.; Weinman, J.; Marteau, T. The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study. *Sociology* **1997**, *31*, 597–606. [[CrossRef](#)]
17. Elliott, R.; Fischer, C.T.; Rennie, D.L. Evolving guidelines for publication of qualitative research studies in psychology and related fields. *Br. J. Clin. Psychol.* **1999**, *27*, 215–229. [[CrossRef](#)]
18. McDonald, N.; Schoenebeck, S.; Forte, A. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* **2019**, *3*, 72. [[CrossRef](#)]
19. Campbell, J.L.; Quincy, C.; Osserman, J.; Pedersen, O.K. Coding In-depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement. *Sociol. Methods Res.* **2013**, *42*, 294–320. [[CrossRef](#)]
20. Corbin, J.; Strauss, A. *Basics of Qualitative Research*, 3rd ed.; Sage: Thousand Oaks, CA, USA, 2008.
21. Patton, M.Q. *Qualitative Research & Evaluation Methods*, 3rd ed.; Sage Publications: Thousand Oaks, CA, USA, 2002.
22. Dunning, D. The Dunning-Kruger Effect: On Being Ignorant of One's Own Ignorance. *Adv. Exp. Soc. Psychol.* **2011**, *44*, 247–296.
23. Collins, P.H. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*; Routledge: Boston, MA, USA, 1990.
24. Du Bois, W.E.B. *The Souls of Black Folk*; Project Gutenberg: Salt Lake City, UT, USA, 1903.
25. Hooks, B. *Feminist Theory: From Margin to Center*; South End Press: Boston, MA, USA, 1984.
26. Cabrera, N.L.; Corces-Zimmerman, C. An Unexamined Life: White Male Racial Ignorance and the Agony of Education for Students of Color. *Equity Excell. Educ.* **2017**, *50*, 300–315. [[CrossRef](#)]
27. Lewis, A.E. "What Group?" Studying Whites and Whiteness in the Era of "Color-Blindness". *Sociol. Theory* **2004**, *22*, 623–646. [[CrossRef](#)]
28. Yosso, T.J.; Smith, W.A.; Ceja, M.; Solórzano, D.G. Critical Race Theory, Racial Microaggressions, and Campus Racial Climate for Latino/a Undergraduates. *Harv. Educ. Rev.* **2009**, *79*, 659–690. [[CrossRef](#)]
29. Winkle-Wagner, R.; McCoy, D.L. Feeling like and "Alien" or "Family"? Comparing students and faculty experiences of diversity in STEM disciplines at a PWI and an HBCU. *Race Ethn. Educ.* **2018**, *21*, 593–606. [[CrossRef](#)]