Intersectional DIF Analyses for a Graduate Student Program Satisfaction Measure

Alejandra Miranda and Michael C. Rodriguez
University of Minnesota

Minnesota Youth Development Research Group
www.mnydrg.com

April 2023

Paper presented at the annual meeting of the
National Council on Measurement in Education, Chicago, IL.

Citation:

Miranda, A., & Rodriguez, M.C. (2023, April 13). *Intersectional DIF analyses for a graduate student program satisfaction measure* [e-board presentation]. National Council on Measurement in Education, Chicago, IL. https://conservancy.umn.edu/handle/11299/194886

Intersectional DIF Analyses for a Graduate Student Program Satisfaction Measure

Abstract

Program satisfaction is important for postsecondary students; therefore, developing psychometrically robust measures is relevant. Using a dataset of graduate students from research universities, DIF analyses by graduate level and international status were performed and compared to DIF analysis crossing level and international status, as finer groups that are relevant to research institutions might provide more meaningful results.

Introduction

College student satisfaction refers to "the contentment or happiness with the total collegial experience, assuming that students, as valued and worthwhile human beings, know what is satisfying or need-fulfilling" (Hallenbeck, 1978, p. 20). It plays an important role in the educational process and is associated with self-efficacy, retention, and performance (DeWitz et al., 2002; Oja, 2011; Starr et al., 1972).

Given the relevance and potential importance of student satisfaction, we need psychometrically sound measures with fair and valid score interpretation. Validity evidence is crucial to support score interpretations for intended uses, as well as fairness, as test-taker responses may be sensitive to individual characteristics (AERA, APA, & NCME, 2014).

To provide evidence about consistency in score interpretation across diverse groups, differential item functioning (DIF) analyses are often conducted. DIF is present when individuals from two or more groups with the same trait level have different expected item responses, which may lead to unfair results and interpretations (Kamata & Vaughn, 2004). However, researchers tend to use broad characteristics such as sex, language, race/ethnicity, and socioeconomic status, ignoring within-group heterogeneity, for example, national origin groups within a given race group; or intersectionality of groups, for example, race by economic-status (Ercikan & Oliveri, 2013).

We performed separate DIF analyses of a program satisfaction measure by graduate student level (master's vs. doctoral) and by international status (international vs domestic). These results were compared to the results from an intersectional DIF analysis crossing graduate level and international status, as a finer definition of groups might lead to more meaningful and relevant DIF results (Ercikan & Oliveri, 2013; Zhang et al., 2005).

Research Question

Do DIF analyses of a measure of program satisfaction by graduate student level and international status identify similar items compared to intersectional DIF analyses by combinations of graduate student level and international status?

Methods

Dataset

To answer the research question, the Graduate Student Experience in the Research University (gradSERU, 2022) dataset was used, developed through a partnership between the Center for Studies in Higher Education at the University of California Berkely and the University of Minnesota. GradSERU was designed to inform institutions about the experiences of graduate students at top-tier research-intensive universities. The gradSERU was administered online in 2021 to graduate students across 15 North American research-intensive universities with over 25,000 participants, to learn about their experiences with peers, faculty, advisor, classes, etc.

Measure

The program satisfaction measure included 13,515 students (age $M = 30.6$ years, $SD = 7.2$ years); 68% doctoral and 32% master's students; 32% international and 68% domestic students.

The program satisfaction measure consisted on 15 items with 4-point rating scales, from *Very dissatisfied* to *Very satisfied*, regarding financial support, quality of the facilities, opportunities for research, knowledge gained, and other aspects of graduate education programs.

Analyses

To measure DIF, Winsteps (Linacre, 2016) was used to estimate the partial credit model and DIF contrasts, based on pairwise differences in item Rasch locations (i.e., item difficulties). DIF levels defined by Educational Testing Service (Linacre, 2016, p. 433) were used to report results where:

- A = negligible DIF: $|\text{DIF}| < 0.43$
- B = slight to moderate DIF: $0.43 \leq |\text{DIF}| < 0.64$
- C = moderate to large DIF: $|\text{DIF}| \geq 0.64$.

To estimate intersectional DIF, four groups of graduate students were compared: master's international (n = 935), master's domestic ($n = 2996$), doctoral international ($n = 3241$), and doctoral domestic ($n = 5880$).

Findings

Measure Quality

Based on confirmatory factor analysis results, there was modest fit overall (RMSEA = .16, CFI = .91, TLI = .90); the CFI and TLI results were adequate, the RMSEA result was not supportive. Given the brief nature of the measure of graduate student program satisfaction, these results are sufficient to pursue DIF analysis, where DIF across relevant groups of students potentially provides one explanation of the modest fit.

DIF Analysis by One Characteristic at a Time

DIF analysis by graduate student level resulted in one flagged item: *satisfaction with departmental climate* (item 1). This item had the lowest value of B-level DIF (contrast of 0.43), with a lower item-location for master's students (measure = 0.29) relative to doctoral students (measure = 0.72). This indicated a higher level of program satisfaction was required for master's students to rate this item similarly relative to doctoral students. There were no C-level DIF items based on graduate student level. No DIF resulted from the analysis by international status.

Intersectional DIF Analysis

The intersectional DIF analysis, graduate student level by international status, flagged three additional items (Table 1). The same item, *satisfaction with departmental climate* (item 1), had B-level DIF (contrast 0.51) with higher item-location for doctoral domestic (0.86) relative to master's domestic student (0.35). This item had B-level DIF (contrast 0.44) between doctoral domestic (0.86) and doctoral international (0.42). The item also had C-level DIF (contrast 0.79) between doctoral domestic (0.86) and master's international student (0.07).

The item regarding *quality of nonacademic student support services* (item 2) had B-level DIF (contrast -0.44) between doctoral international (0.30) and master's international students (-0.14). A higher level of program satisfaction was required for doctoral international students relative to master's international students. This item also had B-level DIF (contrast 0.51) doctoral domestic (0.37) and master's international students (-0.14).

Similarly, the item, *value of the education you are getting for the money you are paying* (item 3), had B-level DIF (contrast 0.47) between doctoral international (0.00) and master's international students (0.47), and also between doctoral domestic (-0.04) and master's international students (0.47; a contrast of -0.52). Finally, the item, *opportunities for research* (item 4), had B-level DIF (-0.44) between doctoral domestic (-0.44) and master's domestic students (0.00). and B-level DIF (-0.57) between doctoral domestic (-0.44) and master's international students (0.13).

Overall, little DIF was found in the items of the program satisfaction measure; only one C-level DIF item was found when comparing doctoral domestic and master's international students, which was not detected in independent DIF analysis.

Intersectional DIF might create more meaningful and useful results. In this study, we demonstrated how using heterogeneous groups might under identify DIF, and how the interpretation of some items in the satisfaction measure may differ for students based on graduate student level and international status.

Table 1
Intersectional DIF Results for Overall Satisfaction Items

Group	Item location	Group	Item location	DIF contrast	Item
DD	0.86	DI	0.42	0.44	Item 1
DD	0.86	MI	0.07	0.79	Item 1
DD	0.86	MD	0.35	0.51	Item 1
DD	0.37	MI	-0.14	0.51	Item 2
MI	-0.14	DI	0.30	-0.44	Item 2
DD	-0.04	MI	0.47	-0.52	Item 3
MI	0.47	DI	0.00	0.47	Item 3
DD	-0.44	MI	0.13	-0.57	Item 4
DD	-0.44	MD	0.00	-0.44	Item 4

Note. DD = doctoral domestic, DI = doctoral international, MD = master's domestic, and MI = master's international.

Conclusions

As students identify with multiple, often complex groups, to understand item functioning in diverse communities requires us to consider intersectional DIF for policy and practice-relevant measures. As we promote the use of the gradSERU, to continually improve graduate education programs and student experiences, such analyses will be important if graduate student experience measures are to be used fairly and validly across groups.

Although there was only one item with C-level DIF, most differences found concerned doctoral domestic students relative to master's international students—perhaps the nature of program satisfaction differs in ways that require additional examination of the construct of program satisfaction. Nevertheless, these results provide support for the use of a measure of program satisfaction meaningfully across these groups of students.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA. https://www.testingstandards.net/

DeWitz, S., & Walsh, W. (2002). Self-efficacy and college student satisfaction. *Journal of Career Assessment, 10*(3), 315-326.

Ercikan, K., & Oliveri, M. E. (2013). Is fairness research doing justice? A modest proposal for an alternative validation approach in differential item functioning (DIF) investigations. In M. Chatterji (Ed.), *Validity, fairness and testing of individuals in high stakes decision-making context* (pp. 69–86). Emerald Publishing.

gradSERU. (2022). *Student Experience in the Research University Survey*. University of Minnesota. https://seru.umn.edu/grad

Hallenbeck, T. (1978). College student satisfaction. *NASPA Journal, 16*(2), 19-25.

Kamata, A., & Vaughn, B. (2004). An introduction to differential item functioning analysis. *Learning Disabilities, 2*(2), 49.

Linacre, J. M. (2016). Winsteps® Rasch measurement computer program user's guide. http://www.winsteps.com/

Oja, M. (2011). Student satisfaction and student performance. *Journa1 of Applied Research in the Community College, 19*(1), 4753-7.

Starr, A., Betz, E., & Menne, J. (1972). Differences in college student satisfaction: Academic dropouts, nonacademic dropouts and nondropouts. *Journal of Counseling Psychology, 19*(4), 318-322.

Zhang, Y., Dorans, N., & Matthews-López, J. (2005). Using DIF dissection method to assess effects of item deletion. *ETS Research Report Series, 2005*(2). https://files.eric.ed.gov/fulltext/ED563094.pdf