

Gender Differences in Career Choice Influences

Tara S. Behrend, Lori Foster Thompson and Adam W. Meade
North Carolina State University

Martha S. Grayson
New York Medical College

Dale A. Newton
East Carolina University

This study examined whether a career influences survey assessing the value medical students place on providing comprehensive patient care exhibited measurement invariance across males and females. Findings supported measurement invariance and indicated that women valued opportunities to provide comprehensive care when choosing a career specialty more than men.

The factors that lead medical students to choose a particular career specialty are complex and poorly understood (Newton, Grayson, & Thompson, 2005). In the past fifteen years, the percentage of medical students choosing primary care careers (e.g. family practice, general internal medicine) has declined sharply, while the percentage of students entering careers such as plastic surgery and emergency medicine has increased dramatically (Newton et al., 2005). Psychiatry and anesthesiology, on the other hand, have fluctuated from year to year (Newton & Grayson, 2003). Many medical professionals believe that unless the decline in students choosing primary care is addressed, the United States will face a shortage of primary care physicians that could be crippling (Schwartz, Basco, Grey, Elmore, & Rubenstein, 2005). For this reason, it is of interest to understand the factors that lead students to select one specialty over another.

Past research has shown that women are more likely than men to choose primary care over other specialties such as anesthesiology (Bland, Meurer, & Maldonado, 1995). This may be a function of gender differences in the values that influence students' career choices. One such value involves the desire to give comprehensive care for one's patients. Comprehensive care entails providing treatment that encompasses psychological and social aspects of patient well-being in addition to biological aspects. Physicians who provide comprehensive patient care are often involved in providing social support to family members, education regarding lifestyle choices, or providing care over a patient's lifespan. As a result, this type of care can be viewed as more

relationship-focused than treatment-focused. Some specialties, including general primary care, are relatively conducive to this type of care, while others, such as urology, are not.

Men's and women's career values are often assessed by asking them to rate the importance they place on various career attributes when choosing a specialty (e.g., Newton et al., 2004). Prior to comparing men's and women's career values, it is important to demonstrate that the survey items used to measure these values function similarly across the two gender groups (Meredith, 1993). In this vein, the goal of the present study is to examine whether men and women differ in how they view and rate survey items designed to assess the importance medical students place on comprehensive patient care when selecting a career.

Mean Differences in Responses to Items Assessing Comprehensive Patient Care

The process of choosing a career is complex and dynamic. One way that students may make career decisions is by assessing the perceived fit between themselves and the career specialty (e.g., personality-career fit; Antony, 1998). This sense of personal fit has been shown to be important for both men and women (Burack et al., 1997). However, perceptions of fit may be dependent on a number of social factors, such as whether a particular specialty is traditionally "male" or "female." Characteristics associated with these specialties may also be weighted as more or less important based on the student's own gender.

Gender influences behaviors and preferences across a variety of contexts. While there is some debate about the degree to which these differences are biologically versus socially influenced, several notable trends have emerged. Among them are gender differences in interdependence and connectedness. For example, Clancy and Dollinger (1993) found that when men and women were asked to select photos that described their lives, women selected more photos of others, while men selected more photos of themselves. This finding supports the notion that women have a tendency to define themselves based on social relationships and connectedness.

Differences in male and female social behavior may also be understood as differences in interdependence (Cross & Madson, 1997; Gabriel & Gardner, 1999). That is, American men tend to view themselves as independent, while American women view themselves as interdependent. This self-construal as either independent or interdependent may form the basis for their cognitions, motivations, and emotions (Cross & Madson, 1997).

The provision of comprehensive patient care involves an interdependent, often long-term, relationship between physician and patient. Further, by understanding the patient's lifestyle, the doctor can get a better and more complete picture of the type of treatment the patient may require. By treating the "whole patient," providers of comprehensive care can offer care that specialists cannot. Given that women are more strongly relationship-oriented than men, and that comprehensive care relies on the formation of a relationship with the patient, it is likely that women will value providing comprehensive care more than men. Thus, we anticipate the following:

Hypothesis: Compared with men, women will rate comprehensive care as more important to their career decisions.

Measurement Invariance of Responses to Items Assessing Comprehensive Patient Care

It is possible that women and men not only place different amounts of value on social relationships, but that they view and define social relationships differently. In a classic study of men and women's jealousy reactions to relationship infidelity, Buss, Larsen, Westen, and Semmelroth (1992) found that while men reacted most strongly to physical infidelity, women's strongest reactions occurred in the case of emotional infidelity. This suggests that women may view and define their relationships differently than men.

Folk wisdom has long noted that women tend to put more effort into social relationships,

manifested in longer phone conversations and social visits. This was confirmed by Roter, Hall, and Aoki (2002), who reported meta-analytic findings demonstrating that female physicians are more likely to engage in interpersonal behaviors, psychosocial question asking, and emotionally focused talk. Further, they found that visits with female physicians lasted an average of two minutes (10%) longer than visits with male physicians. Thus, the type of interpersonal behaviors that men and women consider typical may differ.

Because men and women tend to approach relationships somewhat differently, it is necessary to determine whether they interpret survey questions about relationships in a similar manner. Suppose both men and women were asked to rate the degree to which they value "people skills," for example. Women may view this term positively—people skills could connote interacting in a positive way with others, including forming relationships. In contrast, men could view this term negatively, as connoting a manipulative manner of interaction in order to accomplish a goal. Similarly, the term "fostering relationships" can be viewed either positively or negatively. While men may view the word "relationship" as unprofessional, women could view relationships as crucial to interpersonal interactions. These are only hypothetical examples, but the point is that it is possible for gender-related variables to affect the way survey items are interpreted.

The possibility that men and women not only place different values on relationships with others but also view them differently creates the need for research examining whether survey items measuring relationship-oriented career values, such as comprehensive patient care, are interpreted the same way by men and women. If men and women do not interpret the scale items in the same way, then those items are said to lack measurement invariance (i.e., measurement equivalence). Measurement invariance is the extent to which a scale exhibits the same psychometric properties across groups (or time periods, administration formats, etc.; Horn & McArdle, 1992). If a scale functions differently across groups, observed scale scores cannot be directly interpreted as being indicative of a difference in the latent construct across groups (Horn & McArdle, 1992; Meredith, 1993; Raju, Laffitte, & Byrne, 2002). As such, measurement invariance must be established before it is appropriate to compare mean differences in responses to items assessing comprehensive patient care.

Item Response Theory for Invariance Tests

While confirmatory factor analytic (CFA) methods have commonly been used to assess

measurement invariance for Likert-type data, recently IRT methods of establishing measurement invariance for these data have gained greater acceptance (Raju et al., 2002). IRT methods of establishing measurement invariance typically have used the nomenclature of differential item functioning (DIF). Originally developed for identifying biased test items in dichotomous data, DIF assessments have been adapted to Likert-type scale data. DIF is said to occur when the relationship between levels of examinees' latent trait (θ) and the probability of responses for a particular item differs between two groups (Camilli & Shepard, 1994). IRT is extremely useful for determining how different groups respond to a given scale as IRT provides more information than does CFA, which can lead to a more sensitive test of DIF (Meade & Lautenschlager, 2004; Raju et al., 2002).

In short, a scale used to measure career influences may not function in the same way for women as for men. If an item functions differently for men and women, mean scores cannot be compared across genders (Raju et al., 2002; Taris, Bok, & Meijer, 1998; Vandenberg, 2002). Thus, we used IRT to assess the differential functioning of the measure across gender prior to conducting mean comparisons of importance of comprehensive care.

Research Question: Does a career influences survey assessing the importance of providing comprehensive patient care exhibit measurement invariance across male and female respondents?

Method

Participants

A total of 1363 fourth-year medical students graduating from East Carolina University (35%) and New York Medical College (65%) between 1998 and 2005 completed paper surveys. Approximately 50% of the sample was female ($n = 662$). With regard to ethnicity, 58.8% of the sample was Caucasian, 23.9% was Asian, 7.7% was African American, 1.3% was Hispanic, and 7.6% reported another ethnicity.

Survey

The annual survey examined in this study was administered as part of an on-going research program designed to understand the factors driving medical students' career decisions. The survey contained 98 items, which included questions concerning personal and career-related variables. Early in the survey, respondents were asked to identify their gender. They were also asked to specify which of 23 medical specialties they intended to pursue or select an "other" option. Afterwards, multiple items prompted them to rate the importance

of various job features during the selection of their intended career specialties (e.g., "Emphasizes people skills rather than technical skills"). Nine of the survey items were designed to measure the influence of comprehensive patient care (see Table 1). Respondents were asked to indicate the degree to which these issues pertaining to comprehensive patient care influenced their career choice using a four-point scale: 1 = No influence, 2 = Minor influence, 3 = Moderate influence, 4 = Major influence.

Analysis

One assumption of IRT is that the items reflect a single construct (i.e., they are unidimensional). In order to assess this condition, an exploratory factor analysis was conducted in order to ensure that the comprehensive patient care scale was unidimensional. Once unidimensionality was established, item parameters were estimated separately for each group using the graded response model (GRM; Samejima, 1969) in the MULTILOG 7.03 (Thissen, 1991) program. In order to evaluate items for DIF, item parameters from each group must be placed onto a common metric. The EQUATE program (Baker, 1995) was used to link item parameters from the male and female groups. Finally, DIF and the scale-level counterpart, differential test functioning (DTF), were assessed using the Differential Functioning of Items and Tests program (DFIT; Raju, 1999). The DFIT framework provides a both a non-compensatory DIF index (NCDIF) as well as a compensatory DIF index (CDIF) at the item level. Note, however, that only the NCDIF index is used to evaluate DIF. DFIT has been demonstrated to perform well with Likert-type data (Flowers, Oshima, & Raju, 1999) and has been used extensively in the past (e.g., Collins, Raju, & Edwards, 2000; Donovan, Drasgow, & Probst, 2000; Ellis & Mead, 2000; Fecteau & Craig, 2001; Maurer, Raju, & Collins, 1998).

Results

The first EFA factor explained 56% of the variance in the items, and the scree plot clearly indicated that one dominant factor was present (see Figure 1). Thus, we proceeded with IRT analyses for measurement invariance.

All (equated) item parameters are presented in Table 2 while CDIF and NCDIF values for all items are presented in Table 3. As recommended by Raju (1999), it was items were identified as displaying significant DIF if the NCDIF index exceeded .096. Based on this cutoff criterion, none of the nine items were found to exhibit DIF. Moreover,

the DTF index was .013, well below the .864 cutoff. Thus, results show that no differences were found with regard to the items' interpretation between men and women.

As measurement invariance across genders was supported, it is appropriate to compare men's and women's observed survey responses in order to test the hypothesis driving this study. We predicted that women would rate comprehensive patient care as relatively more important to their career decisions. To test this hypothesis, we compared the average level of importance for the comprehensive patient care scale for men ($M=2.37$, $SD = 0.80$) and women ($M=2.77$, $SD = 0.77$). An independent samples t-test indicated that this difference was significant ($t(1352) = 9.39$, $p < .001$). This demonstrates that women tend to place more importance on comprehensive patient care than do men, thereby supporting the research hypothesis.

Discussion

The current study provides a foundation for researchers seeking to understand career choice influences among medical students. There is an ongoing concern about whether the future supply of primary care physicians will be adequate to meet impending patient care needs. As valuing comprehensive patient care has been shown to predict the career choice among primary care professions (Grayson, Newton, & Thompson, 2005), understanding the factors that influence this value is essential. The present study begins to address this need and illustrates best-practices related to cross-gender comparisons.

Prior to conducting mean comparisons on observed responses, it is important to establish that the scale under investigation measures a different construct in men and women. Put another way, do men and women think about the importance of providing comprehensive care in the same way? The measurement invariance findings demonstrated in this research suggest any differences that may exist are minor and that they do not significantly influence men's and women's interpretations of items on the survey.

Once measurement invariance was established, we tested our research hypothesis, and found that men and women reported different mean levels of importance regarding comprehensive care. Because the survey was determined to be measurement invariant, this mean difference reflects true differences in importance between men and women. Specifically, we found that women place a higher value on comprehensive patient care than men do when selecting a specialty. This is consistent with

past work (e.g., Cross & Madson, 1997; Gabriel & Gardner, 1999; Roter, et al., 2002) indicating that women are more relationship-oriented than are men. Because primary care specialties are likely to involve more comprehensive patient care than non-primary care specialties, this research accords with past evidence that women are more likely than men to choose primary care over other specialties (Bland et al., 1995).

Limitations and Future Research

This study had several limitations which should be noted. Gender differences are to some degree socialized and culturally influenced (Myers, 2005). As most of the medical students in the sample were born in the U.S., differences in interdependence and comprehensive patient care values across gender may not be representative of a broader, non-U.S. population. Further, the data were collected while the students were in their last year of medical school, raising questions about the degree to which participants were able to accurately reflect on the factors that led them to choose a specialty. Finally, participants may have inadvertently placed more emphasis on values they felt were aligned with their specialty choice. Future research should examine this possibility, as well as the influence of other factors (e.g., gender identity, peer influence) on the specialty choices of medical students.

In addition, future studies should move beyond gender and examine other antecedents of the comprehensive patient care value in medical students. Researchers and practitioners interested in increasing the number of individuals pursuing primary care specialties may also wish investigate the utility of considering this value when selecting and placing both male and female medical school students.

Assessing measurement invariance is often neglected in survey research. It vital that researchers interested in attitudinal variables use survey tools that are appropriate for the groups of interest; i.e., tools that have demonstrated measurement invariance. Many studies focus on differences between men and women (or other subgroups), both in the medical field and in the study of human behavior in general. Prior to testing for mean differences, it must be established that the survey instrument is measuring the same thing in each group.

Previous careers research has examined gender differences across a wide variety of work values, often in the absence of measurement invariance analyses. For example, Marini, Fan, Finley, and Beutel (1996) found gender differences in the importance of altruism, intrinsic, and social rewards in a sample of high school seniors entering the workforce, and no difference in the importance of

extrinsic rewards. Without examining the measurement properties of the tools used in studies such as Marini et al.'s, it is impossible to know if observed differences are truly reflective of differences in values.

The present study provides a unique contribution, not only by being the first to demonstrate that male and female medical students differ in the value they place on comprehensive patient care when selecting a career, but also by providing a model of how and why to test for measurement invariance prior to conducting cross-gender comparisons. Future careers research should incorporate a measurement invariance approach in order to more accurately assess group differences in career values.

Practical Implications

On a final note, this research has implications for practice. The subscale examined in this study demonstrated comparable psychometric properties across male and female respondents. This is important for practitioners who may wish to use the scale in the future (e.g., as a diagnostic tool or for the purposes of selecting students who value comprehensive patient care for certain opportunities or assignments). The results of this study suggest that individuals wishing to use this scale for practical purposes can do so with some confidence that the survey items function similarly for men and women. This provides an assurance that the scale is not "working" better for one particular gender at the expense of the other.

References

- Antony, J. S. (1998). Personality-career fit and freshman medical career aspirations: A test of Holland's theory. *Research in Higher Education, 39*, 679-698.
- Baker, F. B. (1995). EQUATE 2.1: Computer program for equating two metrics in item response theory. Madison: University of Wisconsin, Laboratory of Experimental Design.
- Bland, C. J., Meurer, L., & Maldonado, G. (1995). Determinants of primary care specialty choice: A non-statistical meta-analysis of the literature. *Academic Medicine, 70*, 620-641.
- Burack, J. H., Irby, D. M., Carline, J. D., Ambrozy, D. M., Ellsbury, K. E., & Stritter, F. T. (1997). A study of medical students' specialty-choice pathways: Trying on possible selves. *Academic Medicine, 72*, 534-541.
- Buss, D. M., Larsen, R. J., Westen, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science, 3*, 251-255.
- Clancy, S. M., & Dollinger, S. J. (1993). Photographic depictions of the self: Gender and age differences in social connectedness. *Sex Roles, 29*, 477-495.
- Collins, W. C., Raju, N. S., & Edwards, J. E. (2000). Assessing differential functioning in a satisfaction scale. *Journal of Applied Psychology, 85*, 451-461.
- Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*, 5-37.
- Donovan, M. A., Drasgow, F., & Probst, T. M. (2000). Does computerizing paper-and-pencil job attitude scales make a difference? New IRT analyses offer insight. *Journal of Applied Psychology, 85*, 305-13.
- Ellis, B. B., & Mead, A. D. (2000). Assessment of the measurement equivalence of a Spanish translation of the 16PF questionnaire. *Educational and Psychological Measurement, 60*, 787-807.
- Facteau, J. D., & Craig, S. B. (2001). Are performance appraisal ratings from different rating sources comparable? *Journal of Applied Psychology, 86*, 215-27.
- Flowers, C. P., Oshima, T. C., & Raju, N. S. (1999). A description and demonstration of the polytomous-DFIT framework. *Applied Psychological Measurement, 23*, 309-326.
- Gabriel, S., & Gardner, W. L. (1999). Are there "his" and "hers" types of interdependence? The implications of gender differences in collective versus relational interdependence for affect, behavior, and cognition. *Journal of Personality and Social Psychology, 77*, 642-55.
- Grayson, M. S., Newton, D. A., & Thompson, L. F. (2005, November). *Comparing internal medicine to pediatrics: Do the same student-rated career values distinguish generalists from subspecialists in both fields?* Poster presented at the 44th Research in Medical Education segment of the 116th annual meeting of the Association of American Medical Colleges (AAMC), Washington, DC.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance

- in aging research. *Experimental Aging Research*, 18, 117-144.
- Marini, M. M., Fan, P. L., Finley, E., & Beutel, A. M. (1996). Gender and job values. *Sociology of Education*, 69, 49-65.
- Maurer, J., Raju, N. S., & Collins, W. C. (1998). Peer and subordinate performance appraisal measurement equivalence. *Journal of Applied Psychology*, 83, 693-702.
- Meade, A. W., Lautenschlager, G. J., & Johnson, E. C. (in press). A monte-carlo examination of the sensitivity of the DFIT framework for tests of measurement invariance with Likert data. *Applied Psychological Measurement*.
- Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7, 361-388.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.
- Myers, D. G. (2005). *Social psychology* (8th ed.). Boston: McGraw-Hill.
- Newton, D. A., Grayson, M. S., & Thompson, L. F. (2005). The variable influence of lifestyle and income on medical students' career specialty choices: Data from two U.S. medical schools, 1998-2004. *Academic Medicine*, 80, 809-814.
- Newton, D. A., & Grayson, M. S. (2003). Trends in career choice by US medical school graduates. *Journal of the American Medical Association*, 290(9).
- Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, 87, 517-529.
- Raju, N. S., van der Linden, W. J., & Fleer, P. F. (1995). An IRT-based internal measure of test bias with applications for differential item functioning. *Applied Psychological Measurement*, 19, 353-368.
- Roter, D. L., Hall, J. A., & Aoki, Y. (2002). Physician gender effects in medical communication: A meta-analytic review. *The Journal of the American Medical Association*, 288, 756-764.
- Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores* (Psychometric Monograph No. 18). Iowa City, IA: Psychometric Society.
- Schwartz, M., Basco, W. T., Grey, M.R., Elmore, J., & Rubenstein, T. (2005). Rekindling student interest in generalist careers. *Annals of Internal Medicine*, 142, 715-724.
- Thissen, D. (1991). *MULTILOG user's guide* (Version 6.0). Mooresville, IN: Scientific Software.

Author Contact Info:

Tara S. Behrend

Department of Psychology
North Carolina State University
Campus Box 7650
Raleigh, NC 27695-7650
Phone: 919-515-2251
Fax: 919-515-1716
E-mail: tara.behrend@gmail.com

Lori Foster Thompson

Department of Psychology
North Carolina State University
Campus Box 7650
Raleigh, NC 27695-7650
Phone: 919-513-7845
Fax: 919-515-1716
E-mail: lfthompson@ncsu.edu

Adam W. Meade

Department of Psychology
North Carolina State University
Campus Box 7650
Raleigh, NC 27695-7650
Phone: 919-513-4857
Fax: 919-515-1716
E-mail: awmeade@ncsu.edu

Martha S. Grayson

Department of Medicine
New York Medical College
Munger Pavilion, Suite 600
Valhalla, N.Y. 10595
Phone: 914-594-4609
E-mail: grayson@nymc.edu

Dale A. Newton

Department of Pediatrics
3E-139 Brody School of Medicine
East Carolina University
Greenville, NC 27834
phone: 252-744-4422
Fax: 255-744-2398
E-mail: newtond@ecu.edu

Table 1

Items and Communalities

	Comprehensive Patient Care Item	Communalities
1	Includes responsibilities for total care of patient	.547
2	Provides opportunities to manage chronic long-term medical problems	.429
3	Emphasizes people skills rather than technical skills	.382
4	Provides opportunity to be the overall manager of patient care	.499
5	Provides opportunities to practice preventive medicine	.640
6	Allow opportunities to utilize alternative forms of medical therapy	.273
7	Allows development of long-term relationships with patients	.660
8	Emphasizes education of patients for healthier lifestyles	.639
9	Provides opportunities to know patients' families	.623

Table 2

Item Parameters

Item	<i>a</i>		<i>b1</i>		<i>b2</i>		<i>b3</i>	
	Women	Men	Women	Men	Women	Men	Women	Men
1	2.044	2.048	-1.422	-1.442	-.597	-.503	.376	.473
2	1.620	1.993	-.828	-.896	.112	-.066	1.236	.994
3	1.484	1.637	-1.654	-1.715	-.437	-.402	.767	.807
4	1.828	1.842	-1.417	-1.654	-.365	-.471	.782	.833
5	2.503	3.052	-1.189	-1.200	-.481	-.396	.317	.506
6	1.164	1.175	-.693	-.781	.758	.788	2.033	2.077
7	3.308	3.195	-1.464	-1.450	-.876	-.856	-.181	-.082
8	2.828	2.835	-1.532	-1.585	-.689	-.673	.222	.444
9	2.659	3.055	-1.253	-1.387	-.446	-.500	.383	.375

Table 3

DIF Statistics

Item	CDIF	NCDIF
1	.006	.004
2	-.014	.018
3	.000	.000
4	-.003	.005
5	.009	.009
6	.000	.000
7	.007	.004
8	.012	.013
9	-.004	.003

Note: C-DIF = Compensatory DIF index; NC-DIF = Non-compensatory DIF index.

Figure 1. Scree Plot from Exploratory Factor Analysis

