

The Influence of Ethnicity on Patient Satisfaction in a Standardized Patient Assessment

MARTA VAN ZANTEN, JOHN R. BOULET, and DANETTE W. MCKINLEY

ABSTRACT

Purpose. To study possible differences in patient satisfaction ratings as a function of physician and patient ethnicity in a standardized patient (SP) performance-based assessment.

Method. Data from 334,397 ECFMG[®] Clinical Skills Assessment (CSA[®]) simulated clinical encounters were analyzed. A between-groups analysis of covariance was conducted, with independent variables consisting of SP and candidate ethnicity.

Results. Although a significant interaction between SP and candidate ethnicity was found, averaged over all encounters the SPs were equally satisfied with doctors from all cohorts. In general, SPs provided higher satisfaction ratings for racially concordant pairings. Among the ethnic groups of SPs, Asians were the least satisfied.

Conclusions. Satisfaction can vary as a function of the ethnicity of the SP. These results are similar to studies of actual patient encounters that show Asians as less satisfied than patients from other ethnic backgrounds. However, for assessment purposes, provided there is a reasonable mix of SPs, the effect is quite small.

Background

Patient satisfaction is an important component of the physician-patient relationship. Patients who are satisfied with the interaction they have had with their physician are more likely to report better outcomes such as enhanced understanding and adherence to medical regimens.¹⁻⁴ Numerous studies have shown that patients base their level of satisfaction more on physician behaviors and qualities of interaction than on wait times or perceived technical skill.⁵ Some specific physician behaviors that increase patient satisfaction include asking about and discussing psychosocial issues,⁶ talking about treatment effects, and conducting a physical examination.⁷

In addition to physician behaviors, patient satisfaction also appears to be influenced by various physician and patient characteristics, such as gender, age, and ethnicity. Surveys of patients have shown that racial and ethnic minority populations in the United States are typically less satisfied with their health care experiences than the majority population of white Americans.⁸⁻¹³ Among the prevalent minority groups in the United States, Asians^{8,10-12} and linguistic minorities^{9,13} report being the least satisfied with medical care.

Studies have also demonstrated that the ethnicity of a physician can have an influence on a patient's satisfaction with medical care received.¹⁴⁻¹⁹ Saha et al.¹⁵ found that patients with physicians of their own race were more likely to rate their physicians as excellent and report being satisfied with health care received than patients with physicians of a different race. Cooper-Patrick et al.¹⁴ reported similar findings, with patients seeing physicians of their own race being more satisfied and rating their physician's decision-making style as more participatory than patients in race-discordant relationships. Other investigations have concluded that minority patients are more satisfied with race-concordant relationships and deliberately choose physicians of their own race because of personal

preference and language issues, not solely because minority physicians are more likely to be located in their communities.¹⁶⁻²⁰

Within the context of assessing the clinical skills of physicians, physician and patient characteristics and their association with patient satisfaction has been a continuing concern and presents some unique questions. Based on the literature covering actual doctor-patient relationships, one might expect that the interaction between examinee (physician) and patient characteristics could influence scores. Unfortunately, for simulated medical encounters, especially those used for high-stakes assessment purposes, little research has been conducted to specifically explore these potential effects. However, if the scores can be influenced by differences between examiner and candidate, such as ethnicity, then the estimation of candidate abilities may be biased, potentially undermining the validity of the examination.²¹

In the context of a simulated medical environment, the literature describing the influence of race and ethnicity on the doctor-patient relationship supports the theory that standardized patients (SPs) can be trained to provide unbiased, fair ratings. Colliver et al.²² studied the interaction between white and black examinees and SPs and found weak effects and no consistent pattern. Similar results, indicating no evidence of interaction of examinee's and SP's race, were found in a follow-up study that also included Asian examinees.²³ In a study specifically focused on an obstetrics-gynecology clerkship, the authors found no bias attributable to student race in any subsection of interpersonal skills assessment, including an evaluation of patient satisfaction.²⁴

Purpose

The ECFMG is responsible for certifying graduates of international medical schools (IMGs) who wish to pursue graduate medical education in the United States. From July 1998 to May 2004, as part of certification requirements, candidates were required to pass the Clinical Skills Assessment (CSA[®]), a performance-based examination used to ensure that IMGs could demonstrate clinical skills at a level comparable to U.S. medical graduates (USMGs). As part of the assessment, SPs, lay people who were trained to portray common clinical encounters, provided physician satisfaction ratings in each of the simulated encounters.

While there are a number of identifiable SP and physician characteristics that may influence the satisfaction ratings provided to the candidates, the purpose of this study was to look at possible differences in satisfaction ratings as a function of candidate and SP ethnicity. It was hypothesized that, given the simulated environment, the interaction of candidate and SP ethnicity would not significantly impact the ratings given to candidates. Evidence to the contrary, while in accord with most actual doctor-patient-encounter research studies, would suggest that the satisfaction ratings may be influenced by factors not related to the abilities of the candidates.

Method

Measures

The ECFMG[®] CSA is a high-stakes performance-based examination. Physician candidates must demonstrate their clinical skills by interviewing and examining patients in a high-fidelity simulated

environment. Candidates encounter a series of ten or 11 SPs, interacting as they would with actual patients. Candidates have 15 minutes to gather relevant patient data and perform focused physical examinations, as needed, and ten minutes to summarize the data in the form of a clinical note. The test specifications ensure that candidates encounter a varied mix of clinical problems and SPs. Scores for the patient interview, which includes history taking and physical examination, are based on case-specific checklists that are completed by the SPs after the encounter. The clinical notes, completed after the patient encounter, are scored holistically by trained physician raters.

The SPs also evaluate the communication skills of the candidates. Interpersonal skills (IPS) are assessed along four dimensions: interviewing and collecting information, counseling and delivering information, personal manner, and rapport. Spoken-English proficiency is assessed in every encounter using a holistic rating scale. Based on previous CSA research, valid and reliable doctor-patient communication ratings can be obtained from multiple independent SP evaluations.²⁵⁻²⁷

In addition to the scored communication skills ratings, SPs provide a rating of overall satisfaction with the physician. The SPs are trained to base their satisfaction rating on a global sense of likelihood of returning to the physician for future care. The ratings are provided on a five-point scale with the following descriptors: 1 = Definitely no / Overwhelming doubt of examinee's competence and/or interpersonal abilities; 2 = Probably no / Except in case of no other medical care options, would not return; 3 = Neutral / Might return after further consideration, but would seek another physician if easily available; 4 = Probably yes / General impression on this first encounter raised no overwhelming objections; 5 = Definitely yes / Strong sense of competence and interpersonal abilities. Lists of negative feelings (e.g., patronized, worried, confused, annoyed) and positive feelings (e.g., reassured, comforted, comfortable, listened to) are provided to guide the SPs in their ratings. Due to the potential subjective nature of the satisfaction rating, it is not counted as part of a candidate's CSA score and used for research purposes only.

Sample

Data collected in test administrations from July 1998 through December 2003 were analyzed to examine the effects of SP and candidate ethnicity on satisfaction ratings. Ratings from 334,397 encounters were included. The candidate sample consisted of 27,569 IMGs. Of these candidates, 47.4% were Asian, 5.8% were black, 9.1% were Hispanic, and 37.7% were white. Of the analysis cohort, 57.8% were male, and 52.5% were age 30 or younger (mean age = 31.5). These encounters included ratings from a sample of 247 SPs, of which 2.8% were Asian, 42.9% were black, 4.0% were Hispanic, and 50.2% were white. Female SPs comprised 57.5% of those included in these analyses. The cases the SPs portrayed are classified in three age categories: 18-44, 45-64, and over 65. Of the 247 SPs, 45.3% portrayed cases classified as 18-44, 34.8% portrayed cases in the 45-64 category, and 19.8% portrayed cases in the 65-and-over category.

Results

A 4 × 4 between-groups analysis of covariance was conducted on satisfaction ratings. Independent variables consisted of SP and candidate ethnicity (Asian, black, Hispanic, and white). Covariates were interpersonal skills and spoken-English proficiency ratings. These covariates were included to take into account the fact that SPs, over time, encounter candidates of varying ability. After adjustment for candidate ability, a significant interaction between SP and candidate ethnicity was found ($F_{9,334379} = 21.3, p < .01$), with SPs generally providing higher satisfaction ratings for encounters in which there was racial concordance. There was a statistically significant main effect for SP ethnicity ($F_{3,10} = 24.9, p < .01$), with

TABLE 1. Mean Satisfaction Ratings (Standard Deviations) of Standardized Patients and Candidates According to Their Ethnicity

Standardized Patient	Candidate				
	Asian <i>n</i> = 13,076	Black <i>n</i> = 1,586	Hispanic <i>n</i> = 2,510	White <i>n</i> = 10,397	Overall <i>n</i> = 27,569
Asian (<i>n</i> = 7)	3.2 (.9)	3.2 (1.0)	3.2 (.9)	3.3 (1.0)	3.2 (.9)
Black (<i>n</i> = 106)	3.5 (.9)	3.6 (.9)	3.6 (.9)	3.6 (.9)	3.6 (.9)
Hispanic (<i>n</i> = 10)	3.6 (.8)	3.6 (.8)	3.7 (.8)	3.7 (.8)	3.7 (.8)
White (<i>n</i> = 124)	3.4 (.9)	3.5 (.9)	3.5 (.9)	3.6 (.9)	3.5 (.9)
Overall (<i>n</i> = 247)	3.5 (.9)	3.5 (.9)	3.5 (.9)	3.6 (.9)	

Asian SPs providing significantly lower satisfaction ratings, and Hispanic SPs providing the highest ratings. There was no statistically significant main effect due to candidate ethnicity ($F_{3,10} = 1.4, p = .298$).

Mean satisfaction ratings by candidate and SP ethnicity are presented in Table 1. Based on the summary statistics in this table, it is clear that satisfaction ratings are related to the ethnicity of the SP and the ethnicity of the candidate. Although there was a statistically significant SP ethnicity by candidate ethnicity interaction, this result was based on an extremely large sample. Interestingly, for all SP ethnic groups, there was relatively little variation in mean scores as a function of candidate ethnicity. However, Asian SPs provided, on average, the lowest satisfaction ratings (mean = 3.2, SD = .9). Based on the mean, candidates encountering an Asian SP would receive a satisfaction rating that was approximately .4 standard deviations lower than that provided by SPs in any of the other three ethnic cohorts. Nevertheless, all of the candidate groups, including Asians, received similar patient satisfaction ratings. Averaged over all SP encounters, the SPs were equally satisfied with Asian, black, Hispanic, and white doctors.

Discussion

For performance-based SP examinations, especially those with high-stakes consequences, it is imperative that potential measurement errors are controlled. When rating scales are employed, there is always the possibility that scores will be influenced by traits other than those being measured as part of the assessment. The literature would suggest that, with proper SP training and quality assurance procedures, the interaction of SP and examinee characteristics on scores or ratings will be minimal.^{22-24,28} However, as part of a score validation process, it is still important to investigate potential sources of systematic measurement error.

The ECFMG CSA has been used for over five years to assess the clinical skills of IMGs. As part of the test development process, many different types of cases are constructed. To model these cases appropriately, individuals with various characteristics (e.g., gender, age, ethnicity) are required. Although patient ethnicity may be a valid risk factor for certain diseases, and potentially have some effect on the perceived competence of the physician, it should not unduly influence scores. For CSA, SPs are recruited from a number of ethnic groups. Although patient ethnicity is not a specification of the test blueprint, the availability of an ethnically diverse SP pool ensures that examinees encounter, on average, a group of SPs with a heterogeneous mix of racial attributes.

Our results indicate that, after controlling for communication abilities of the candidates, satisfaction with the physician can vary somewhat as a function of the ethnicity of the standardized patient. While black, white, and Hispanic SPs were, on average, equally satisfied with the physicians, Asian SPs provided significantly lower ratings. Interestingly, based on previous health care studies, Asian patients tend to be less satisfied with their care.^{8,10-12} The fact that our results mirror reality could be interpreted as supporting the

validity of the CSA examination. However, if the patients are truly "standardized," then one would not expect satisfaction to vary as a function of SP demographics, be it ethnicity, gender, age, socioeconomic status, or some other attribute. Nevertheless, the overall effect was reasonably small and was based on the ethnic cohort with the fewest members ($n = 7$). Thus, if an individual candidate's test form includes a mix of SP ethnicities, the overall effect would be expected to be negligible.

There are a number of limitations of this study that warrant further discussion. First, the SP effects noted apply to patient satisfaction, a somewhat nebulous construct that is dependent on the individual's expectations of the health care provider. Nevertheless, one might expect that our findings would generalize to other scored parts of the CSA, namely the evaluation of interpersonal skills and spoken-English proficiency. If SP rating tendencies are not controlled, either through training or by some statistical adjustment, certain candidates may be advantaged or disadvantaged. Additional research in this area is certainly needed. Second, our results are based on averaged values and may not reflect individual SP biases. Here, it would be informative to investigate rating patterns across cases where there are two SPs of different ethnicities. Third, our results are based on simulated cases where the interaction is modeled on a first-time visit and the patient has no choice of physician. The results could be quite different in real settings where the patients deliberately chose a health care provider and rapport can be built over time. Fourth, although the SP ratings may have been influenced by the ethnicity of the candidate, this would be based on perception only. The analyses used the self-declared ethnicity of the candidate. To the extent that these two labels differ, the results may change. Fifth, although interpersonal skills and English-language proficiency were used as covariates, a physician's history-taking and physical examination ability was not part of the analysis. An SP's perception of a candidate's knowledge and ability, above and beyond communication skills, would also have been likely to influence patient satisfaction ratings. Finally, the results of this study are generalizable only to graduates of international medical schools. The satisfaction of SPs interacting with graduates of U.S. medical schools, regardless of ethnicity, may be different due to greater cultural similarities.

Although standardized patients undergo extensive training, it is reasonable to expect some variation in their ratings of physician satisfaction. While SPs of one ethnicity tended to be more satisfied with physicians of the same ethnicity, the effects at the encounter level were reasonably small and, based on a multistation assessment, would tend to cancel out. Additional potential sources of score invalidity, including other patient and physician characteristics, should be investigated.

Correspondence: Marta van Zanten, ECFMG, 3624 Market Street, Philadelphia, PA, 19104; e-mail: (mvanzanten@ecfm.org).

References

- Ong LM, de Haes JC, Hoos AM, Lammes FB. Doctor-patient communication: a review of the literature. *Soc Sci Med.* 1995;40:903-18.
- Williams S, Weinman J, Dale J. Doctor-patient communication and patient satisfaction: a review. *Fam Pract.* 1998;15:480-92.
- Rowland-Morin PA, Carroll JG. Verbal communication skills and patient satisfaction. A study of doctor-patient interviews. *Eval Health Prof.* 1990;13:168-85.
- Beck RS, Daughtridge R, Sloane PD. Physician-patient communication in the primary care office: a systematic review. *J Am Board Fam Pract.* 2002;15:25-38.
- Boudreaux ED, O'Hea EL. Patient satisfaction in the emergency department: a review of the literature and implications for practice. *J Emerg Med.* 2004;26:13-26.
- Bertakis KD, Roter D, Putnam SM. The relationship of physician medical interview style to patient satisfaction. *J Fam Pract.* 1991;32:175-81.
- Robbins JA, Bertakis KD, Helms LJ, Azari R, Callahan EJ, Creten DA. The influence of physician practice behaviors on patient satisfaction. *Fam Med.* 1993; 25:17-20.
- Murray-Garcia JL, Selby JV, Schmittiel J, Grumbach K, Quesenberry CP Jr. Racial and ethnic differences in a patient survey: patients' values, ratings, and reports regarding physician primary care performance in a large health maintenance organization. *Med Care.* 2000;38:300-10.
- Weech-Maldonado R, Morales LS, Elliott M, Spritzer K, Marshall G, Hays RD. Race/ethnicity, language, and patients' assessments of care in Medicaid managed care. *Health Serv Res.* 2003;38:789-808.
- Haviland MG, Morales LS, Reise SP, Hays RD. Do health care ratings differ by race or ethnicity? *Jt Comm J Qual Saf.* 2003;29:134-45.
- Taira DA, Safran DG, Seto TB, et al. Asian-American patient ratings of physician primary care performance. *J Gen Intern Med.* 1997;12:237-42.
- Taira DA, Safran DG, Seto TB, et al. Do patient assessments of primary care differ by patient ethnicity? *Health Serv Res.* 2001;36(6 Pt 1):1059-71.
- Morales LS, Cunningham WE, Brown JA, Liu H, Hays RD. Are Latinos less satisfied with communication by health care providers? *J Gen Intern Med.* 1999;14:409-17.
- Cooper-Patrick L, Gallo JJ, Gonzales JJ, et al. Race, gender, and partnership in the patient-physician relationship. *JAMA.* 1999;282:583-9.
- Saha S, Komaromy M, Koepsell TD, Bindman AB. Patient-physical racial concordance and the perceived quality and use of health care. *Arch Intern Med.* 2002;159:997-1004.
- Gray B, Stoddard JJ. Patient-physician pairing: does racial and ethnic congruity influence selection of a regular physician? *J Commun Health.* 2002;22:247-59.
- Saha S, Taggart SH, Komaromy M, Bindman AB. Do patients choose physicians of their own race? *Minority Health.* 2003;19:78-83.
- Laveist TA, Nuru-Jeter A. Is doctor-patient race concordance associated with greater satisfaction with care? *J Health Soc Behav.* 2002;43:296-306.
- Laveist TA, Carroll T. Race of physician and satisfaction with care among African-American patients. *J Natl Med Assoc.* 2002;94:937-43.
- Cooper LA, Roter DL, Johnson RL, Ford DE, Steinwachs DM, Powe NR. Patient-centered communication, ratings of care, and concordance of patient and physician race. *Ann Intern Med.* 2003;139:907-15.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association, 1999.
- Colliver JA, Swartz MH, Robbs RS. The effect of examinee and patient ethnicity in clinical-skills assessment with standardized patients. *Adv Health Sci Educ.* 2001;6:5-13.
- Swartz MH, Colliver JA, Robbs RS. The interaction of examinee's ethnicity and standardized patient's ethnicity: an extended analysis. *Acad Med.* 2001;76(10 suppl):S96-8.
- Bienstock JL, Tzou WS, Martin SA, Fox HE. Effect of student ethnicity on interpersonal skills and objective standardized clinical examination scores. *Obstet Gynecol.* 2000;96:1011-3.
- Friedman M, Sutnick AI, Stillman PL, et al. The use of standardized patients to evaluate the spoken-English proficiency of foreign medical graduates. *Acad Med.* 1991;66(suppl 9):S61-3.
- Boulet JR, Ben David MF, et al. Using standardized patients to assess the interpersonal skills of physicians. *Acad Med.* 1998;73(10 suppl):S94-6.
- Boulet JR, van Zanten M, McKinley DW, Gary NE. Evaluating the spoken English proficiency of graduates of foreign medical schools. *Med Educ.* 2001;35:767-73.
- van Zanten M, Boulet JR, McKinley D, Whelan GP. Evaluating the spoken English proficiency of international medical graduates: detecting threats to the validity of standardized patient ratings. *Med Educ.* 2003;37:69-76.