MICHAEL R.N. MCDONNELL INNS OF COURT

PRESENTATION TEAM #1
NOVEMBER 14, 2017

Who? Us Biased??

MICHAEL R.N. MCDONNELL
AMERICAN INNS OF COURT
GROUP I PRESENTATION
November 13, 2017

Who?? Us Biased?

LEARNING OBJECTIVES:

As a result of attending this presentation, attendees will be able to:

a) Better understand and address the unconscious, systematic ways in which we receive and categorize information and people.

b) Consider how implicit bias may affect our legal practice, advocacy and decision making.

c) Create a culture and environment within our profession for accountability to address bias, implicit or otherwise, that is devoid of defensiveness and conducive to productive discussion and solutions.

d) Identify and understand how the Florida Bar Rules of Professional Conduct (Chapter 4) Rules 4-1 through 4-8, Rules, the Canons of the Florida Code of Judicial Conduct (3)(b)(5)-(6) may apply to our consideration of bias, implicit and otherwise.

111 So.3d 206
District Court of Appeal of Florida,
Fourth District.

Glenda MARTINEZ, Petitioner,
v.
John CRAMER and J. Alan Smith, Respondents.

No. 4D13–368.
|
April 3, 2013.

Synopsis
Background: Guardian filed petition for writ of prohibition seeking to prevent trial judge from presiding over further proceedings in the Circuit Court, Fifteenth Judicial Circuit, Palm Beach County, Martin H. Colin, J., in guardianship of the guardian's spouse, for whom she was the designated healthcare surrogate.

[Holding:] The District Court of Appeal held that trial judge's personal comments and actions toward guardian warranted disqualification.

Petition granted.

West Headnotes (5)

[1] **Judges**
Sufficiency of objection, affidavit, or motion

The substantive test for whether a motion to disqualify a judge is legally sufficient is whether the facts alleged would place a reasonably prudent person in fear of not receiving a fair and impartial trial.

Cases that cite this headnote

[2] **Judges**
Bias and Prejudice

When considering a motion to disqualify, it is not a question of what the judge feels, but the feeling in the mind of the party seeking to disqualify and the basis for that feeling.

Cases that cite this headnote

[3] **Judges**
Bias and Prejudice

A litigant's fear of judicial bias must be objectively reasonable in order to warrant disqualification.

Cases that cite this headnote

[4] **Judges**
Bias and Prejudice

The reasons for disqualification of a judge must show personal bias or prejudice to the litigant in order to warrant disqualification.

Cases that cite this headnote

[5] **Judges**
Statements and expressions of opinion by judge

Trial judge's actions in ejecting guardian from the courtroom and striking guardian's testimony on the basis of a perceived insult to judge, and his comment that guardian's entire demeanor, including that "her face, her voice, her sound, may have been unpleasant to everyone else," save the ward, would have lead any reasonably prudent person to fear that she would not receive a fair hearing before the judge, and therefore warranted disqualification in guardianship proceeding; judge's actions as well as his personal comments about guardian, when taken

in combination, far exceeded comments or actions necessary to control his courtroom and were sufficient to evidence to a reasonable person bias requiring disqualification, even if the judge may have felt that he had no bias.

1 Cases that cite this headnote

Attorneys and Law Firms

***207** Robin Bresky of Law Offices of Robin Bresky, Boca Raton, for petitioner.

Ellen S. Morris of Elder Law Associates PA, Boca Raton, for respondent John Cramer.

Opinion

PER CURIAM.

Petitioner, Glenda Martinez, petitions for writ of prohibition to prevent the trial judge from presiding over further proceedings in the guardianship of the petitioner's spouse, for whom she is the designated healthcare surrogate. We grant the petition.

[1] [2] [3] [4] The substantive test for whether a motion to disqualify is legally sufficient is "whether the facts alleged would place a reasonably prudent person in fear of not receiving a fair and impartial trial." *Santisteban v. State,* 72 So.3d 187, 193 (Fla. 4th DCA 2011). "It is not a question of what the judge feels, but the feeling in the mind of the party seeking to disqualify and the basis for that feeling." *Corie v. City of Riviera Beach,* 954 So.2d 68, 70 (Fla. 4th DCA 2007). It must, however, be objectively reasonable. *Foy v. State,* 818 So.2d 704, 706 (Fla. 5th DCA 2002). The reasons for disqualification must also show personal bias or prejudice to the litigant. *Levine v. State,* 650 So.2d 666, 667 (Fla. 4th DCA 1995).

[5] We have reviewed the verified motion and the transcript of proceedings on which the motion was based. We conclude that the judge's acts of ejecting petitioner from the courtroom, later striking her testimony on the basis of a perceived insult to him, and his comment that petitioner's entire demeanor, including that "her face, her voice, her sound, maybe unpleasant to everyone else," save the ward, would lead any reasonably prudent person to fear that she would not receive a fair hearing before the judge. *See, e.g., Colarusso v. Colarusso,* 20 So.3d 985, 986 (Fla. 3d DCA 2009).

Although respondent contends that a judge's comments in the course of attempting to control the courtroom are ordinarily not legally sufficient to require disqualification, *see Braddy v. State,* 111 So.3d 810, 833–34, 2012 WL 5514368 (Fla. Nov. 15, 2012), we conclude that the judge's actions as well as his personal comments about petitioner, when taken in combination, far exceeded comments or actions necessary to control his courtroom and were sufficient to evidence to a reasonable person bias requiring disqualification, even if the judge may have felt that he had no bias.

Therefore, we grant the petition but withhold the formal issuance of the writ, as we are sure that the trial judge will act in conformance with this opinion and remit his disqualification so that the matter may be assigned to a different judge.

MAY, C.J., WARNER and CONNER, JJ., concur.

All Citations

111 So.3d 206, 38 Fla. L. Weekly D744

End of Document


Opinions
Supreme Court Opinions
1st DCA Opinions
2nd DCA Opinions
3rd DCA Opinions
4th DCA Opinions
5th DCA Opinions
Unofficial Archive of Opinions

Judicial Ethics
Code of Conduct
Canon 7 Guide
JEAC Opinions

Court Rules
Florida Rules Of Procedure
Amendments to Approved Rules
Proposed Rules
Rules for Mediators

Court Orders
Clerk's Office

Other Supreme Court Documents
Other Court Rules Resources

Opinions Banner Image

Code of Judicial Conduct

Canon 3

A Judge Shall Perform the Duties of Judicial Office Impartially and Diligently

A. Judicial Duties in General.

The judicial duties of a judge take precedence over all the judge's other activities. The judge's judicial duties include all the duties of the judge's office prescribed by law. In the performance of these duties, the specific standards set forth in the following sections apply.

B. Adjudicative Responsibilities.

(1) A judge shall hear and decide matters assigned to the judge except those in which disqualification is required.

(2) A judge shall be faithful to the law and maintain professional competence in it. A judge shall not be swayed by partisan interests, public clamor, or fear of criticism.

(3) A judge shall require order and decorum in proceedings before the judge.

(4) A judge shall be patient, dignified, and courteous to litigants, jurors, witnesses, lawyers, and others with whom the judge deals in an official capacity, and shall require similar conduct of lawyers, and of staff, court officials, and others subject to the judge's direction and control.

(5) A judge shall perform judicial duties without bias or prejudice. A judge shall not, in the performance of judicial duties, by words or conduct manifest bias or prejudice, including but not limited to bias or prejudice based upon race, sex, religion, national origin, disability, age, sexual orientation, or socioeconomic status, and shall not permit staff, court officials, and others subject to the judge's direction and control to do so. This section does not preclude the consideration of race, sex, religion, national origin, disability, age, sexual orientation, socioeconomic status, or other similar factors when they are issues in the proceeding.

(6) A judge shall require lawyers in proceedings before the judge to refrain from manifesting, by words, gestures, or other conduct, bias or prejudice based upon race, sex, religion, national origin, disability, age, sexual orientation, or socioeconomic status, against parties, witnesses, counsel, or others. This Section 3B(6) does not preclude legitimate advocacy when race, sex, religion, national origin, disability, age, sexual orientation, socioeconomic status, or other similar factors are issues in the proceeding.

(7) A judge shall accord to every person who has a legal interest in a proceeding, or that person's lawyer, the right to be heard according to law. A judge shall not initiate, permit, or consider ex parte communications, or consider other communications made to the judge outside the presence of the parties concerning a pending or impending proceeding except that:

(a) Where circumstances require, ex parte communications for scheduling, administrative purposes, or emergencies that do not deal with substantive matters or issues on the merits are authorized, provided:

(i) the judge reasonably believes that no party will gain a procedural or tactical advantage as a result of the ex parte communication, and

(ii) the judge makes provision promptly to notify all other parties of the substance of the ex parte communication and allows an opportunity to respond.

(b) A judge may obtain the advice of a disinterested expert on the law applicable to a proceeding before the judge if the judge gives notice to the parties of the person consulted and the substance of the advice and affords the parties reasonable opportunity to respond.

(c) A judge may consult with other judges or with court personnel whose function is to aid the judge in carrying out the judge's adjudicative responsibilities.

(d) A judge may, with the consent of the parties, confer separately with the parties and their lawyers in an effort to mediate or settle matters pending before the judge.

(e) A judge may initiate or consider any ex parte communications when expressly authorized by law to do so.

(8) A judge shall dispose of all judicial matters promptly, efficiently, and fairly.

(9) A judge shall not, while a proceeding is pending or impending in any court, make any public comment that might reasonably be expected to affect its outcome or impair its fairness or make any nonpublic comment that might substantially interfere with a fair trial or hearing. The judge shall require similar abstention on the part of court personnel subject to the judge's direction and control. This Section does not prohibit judges from making public statements in the course of their official duties or from explaining for public information the procedures of the court. This Section does not apply to proceedings in which the judge is a litigant in a personal capacity.

(10) A judge shall not, with respect to parties or classes of parties, cases, controversies or issues likely to come before the court, make pledges, promises or commitments that are inconsistent with the impartial performance of the adjudicative duties of the office.

(11) A judge shall not commend or criticize jurors for their verdict other than in a court order or opinion in a proceeding, but may express appreciation to jurors for their service to the judicial system and the community.

(12) A judge shall not disclose or use, for any purpose unrelated to judicial duties, nonpublic information acquired in a judicial capacity.

C. Administrative Responsibilities.

(1) A judge shall diligently discharge the judge's administrative responsibilities without bias or prejudice and maintain professional competence in judicial administration, and should cooperate with other judges and court officials in the administration of court business.

(2) A judge shall require staff, court officials, and others subject to the judge's direction and control to observe the standards of fidelity and diligence that apply to the judge and to refrain from manifesting bias or prejudice in the performance of their official duties.

(3) A judge with supervisory authority for the judicial performance of other judges shall take reasonable measures to assure the prompt disposition of matters before them and the proper performance of their other judicial responsibilities.

(4) A judge shall not make unnecessary appointments. A judge shall exercise the power of appointment impartially and on the basis of merit. A judge shall avoid nepotism and favoritism. A judge shall not approve compensation of appointees beyond the fair value of services rendered.

D. Disciplinary Responsibilities.

(1) A judge who receives information or has actual knowledge that substantial likelihood exists that another judge has committed a violation of this Code shall take appropriate action.

(2) A judge who receives information or has actual knowledge that substantial likelihood exists that a lawyer has committed a violation of the Rules Regulating The Florida Bar shall take appropriate action.

(3) Acts of a judge, in the discharge of disciplinary responsibilities, required or permitted by Sections 3D(1) and 3D(2) are part of a judge's judicial duties and shall be absolutely privileged, and no civil action predicated thereon may be instituted against the judge.

E. Disqualification.

(1) A judge shall disqualify himself or herself in a proceeding in which the judge's impartiality might reasonably be questioned, including but not limited to instances where:

(a) the judge has a personal bias or prejudice concerning a party or a party's lawyer, or personal knowledge of disputed evidentiary facts concerning the proceeding;

(b) the judge served as a lawyer or was the lower court judge in the matter in controversy, or a lawyer with whom the judge previously practiced law served during such association as a lawyer concerning the matter, or the judge has been a material witness concerning it;

(c) the judge knows that he or she individually or as a fiduciary, or the judge's spouse, parent, or child wherever residing, or any other member of the judge's family residing in the judge's household has an economic interest in the subject matter in controversy

or in a party to the proceeding or has any other more than de minimis interest that could be substantially affected by the proceeding;

(d) the judge or the judge's spouse, or a person within the third degree of relationship to either of them, or the spouse of such a person:

(i) is a party to the proceeding, or an officer, director, or trustee of a party;

(ii) is acting as a lawyer in the proceeding;

(iii) is known by the judge to have a more than de minimis interest that could be substantially affected by the proceeding;

(iv) is to the judge's knowledge likely to be a material witness in the proceeding;

(e) the judge's spouse or a person within the third degree of relationship to the judge participated as a lower court judge in a decision to be reviewed by the judge.

(f) the judge, while a judge or a candidate for judicial office, has made a public statement that commits, or appears to commit, the judge with respect to:

(i) parties or classes of parties in the proceeding;

(ii) an issue in the proceeding; or

(iii) the controversy in the proceeding.

(2) A judge should keep informed about the judge's personal and fiduciary economic interests, and make a reasonable effort to keep informed about the economic interests of the judge's spouse and minor children residing in the judge's household.

F. Remittal of Disqualification.

A judge disqualified by the terms of Section 3E may disclose on the record the basis of the judge's disqualification and may ask the parties and their lawyers to consider, out of the presence of the judge, whether to waive disqualification. If following disclosure of any basis for disqualification other than personal bias or prejudice concerning a party, the parties and lawyers, without participation by the judge, all agree the judge should not be disqualified, and the judge is then willing to participate, the judge may participate in the proceeding. The agreement shall be incorporated in the record of the proceeding.

COMMENTARY

Canon 3B(4). The duty to hear all proceedings fairly and with patience is not inconsistent with the duty to dispose promptly of the business of the court. Judges can be efficient and business-like while being patient and deliberate.

Canon 3B(5). A judge must refrain from speech, gestures or other conduct that could reasonably be perceived as sexual harassment and must require the same standard of conduct of others subject to the judge's direction and control.
A judge must perform judicial duties impartially and fairly. A judge who manifests bias on any basis in a proceeding impairs the fairness of the proceeding and brings the judiciary into disrepute. Facial expression and body language, in addition to oral communication, can give to parties or lawyers in the proceeding, jurors, the media and others an appearance of judicial bias. A judge must be alert to avoid behavior that may be perceived as prejudicial.

Canon 3B(7). The proscription against communications concerning a proceeding includes communications from lawyers, law teachers, and other persons who are not participants in the proceeding, except to the limited extent permitted.

To the extent reasonably possible, all parties or their lawyers shall be included in communications with a judge.

Whenever presence of a party or notice to a party is required by Section 3B(7), it is the party's lawyer, or if the party is unrepresented, the party who is to be present or to whom notice is to be given.

An appropriate and often desirable procedure for a court to obtain the advice of a disinterested expert on legal issues is to invite the expert to file a brief as amicus curiae.

Certain ex parte communication is approved by Section 3B(7) to facilitate scheduling and other administrative purposes and to accommodate emergencies. In general, however, a judge must discourage ex parte communication and allow it only if all the criteria stated in Section 3B(7) are clearly met. A judge must disclose to all parties all ex parte communications described in Sections 3B(7)(a) and 3B(7)(b) regarding a proceeding pending or impending before the judge.
A judge must not independently investigate facts in a case and must consider only the evidence presented.

A judge may request a party to submit proposed findings of fact and conclusions of law, so long as the other parties are apprised of the request and are given an opportunity to respond to the proposed findings and conclusions.

A judge must make reasonable efforts, including the provision of appropriate supervision, to ensure that Section 3B(7) is not violated through law clerks or other personnel on the judge's staff.

If communication between the trial judge and the appellate court with respect to a proceeding is permitted, a copy of any written communication or the substance of any oral communication should be provided to all parties.

Canon 3B(8). In disposing of matters promptly, efficiently, and fairly, a judge must demonstrate due regard for the rights of the parties to be heard and to have issues resolved without unnecessary cost or delay. Containing costs while preserving fundamental rights of parties also protects the interests of witnesses and the general public. A judge should monitor and supervise cases so as to reduce or eliminate dilatory practices, avoidable delays, and unnecessary costs. A judge should encourage and seek to facilitate settlement, but parties should not feel coerced into surrendering the right to have their controversy resolved by the courts.

Prompt disposition of the court's business requires a judge to devote adequate time to judicial duties, to be punctual in attending court and expeditious in determining matters under submission, and to insist that court officials, litigants, and their lawyers cooperate with the judge to that end.

Canon 3B(9) and 3B(10). Sections 3B(9) and (10) restrictions on judicial speech are essential to the maintenance of the integrity, impartiality, and independence of the judiciary. A pending proceeding is one that has begun but not yet reached final disposition. An impending proceeding is one that is anticipated but not yet begun. The requirement that judges abstain from public comment regarding a pending or impending proceeding continues during any appellate process and until final disposition. Sections 3B(9) and (10) do not prohibit a judge from commenting on proceedings in which the judge is a litigant in a personal capacity, but in cases such as a writ of mandamus where the judge is a litigant in an official capacity, the judge must not comment publicly. The conduct of lawyers relating to trial publicity is governed by Rule 4-3.6 of the Rules Regulating The Florida Bar.

Canon 3B(10). Commending or criticizing jurors for their verdict may imply a judicial expectation in future cases and may impair a juror's ability to be fair and impartial in a subsequent case.

Canon 3C(4). Appointees of a judge include assigned counsel, officials such as referees, commissioners, special magistrates, receivers, mediators, arbitrators, and guardians and personnel such as clerks, secretaries, and bailiffs. Consent by the parties to an appointment or an award of compensation does not relieve the judge of the obligation prescribed by Section 3C(4). See also Fla.Stat. § 112.3135 (1991).

Canon 3D. Appropriate action may include direct communication with the judge or lawyer who has committed the violation, other direct action if available, or reporting the violation to the appropriate authority or other agency. If the conduct is minor, the Canon allows a judge to address the problem solely by direct communication with the offender. A judge having knowledge, however, that another judge has committed a violation of this Code that raises a substantial question as to that other judge's fitness for office or has knowledge that a lawyer has committed a violation of the Rules of Professional Conduct that raises a substantial question as to the lawyer's honesty, trustworthiness or fitness as a lawyer in other respects, is required under this Canon to inform the appropriate authority. While worded differently, this Code provision has the identical purpose as the related Model Code provisions.

Canon 3E(1). Under this rule, a judge is disqualified whenever the judge's impartiality might reasonably be questioned, regardless of whether any of the specific rules in Section 3E(1) apply. For example, if a judge were in the process of negotiating for employment with a law firm, the judge would be disqualified from any matters in which that law firm appeared, unless the disqualification was waived by the parties after disclosure by the judge.

A judge should disclose on the record information that the judge believes the parties or their lawyers might consider relevant to the question of disqualification, even if the judge believes there is no real basis for disqualification. The fact that the judge conveys this information does not automatically require the judge to be disqualified upon a request by either party, but the issue should be resolved on a case-by-case basis. Similarly, if a lawyer or party has previously filed a complaint against the judge with the Judicial Qualifications Commission, that the fact does not automatically require disqualification of the judge. Such disqualification should be on a case-by-case basis.

By decisional law, the rule of necessity may override the rule of disqualification. For example, a judge might be required to participate in judicial review of a judicial salary statute, or might be the only judge available in a matter requiring immediate judicial action, such as a hearing on probable cause or a temporary restraining order. In the

latter case, the judge must disclose on the record the basis for possible disqualification and use reasonable efforts to transfer the matter to another judge as soon as practicable.

Canon 3E(1)(b). A lawyer in a government agency does not ordinarily have an association with other lawyers employed by that agency within the meaning of Section 3E(1)(b); a judge formerly employed by a government agency, however, should disqualify himself or herself in a proceeding if the judge's impartiality might reasonably be questioned because of such association.

Canon 3E(1)(d). The fact that a lawyer in a proceeding is affiliated with a law firm with which a relative of the judge is affiliated does not of itself disqualify the judge. Under appropriate circumstances, the fact that "the judge's impartiality might reasonably be questioned" under Section 3E(1), or that the relative is known by the judge to have an interest in the law firm that could be "substantially affected by the outcome of the proceeding" under Section 3E(1)(d)(iii) may require the judge's disqualification.

Canon 3E(1)(e). It is not uncommon for a judge's spouse or a person within the third degree of relationship to a judge to also serve as a judge in either the trial or appellate courts. However, where a judge exercises appellate authority over another judge, and that other judge is either a spouse or a relationship within the third degree, then this Code requires disqualification of the judge that is exercising appellate authority. This Code, under these circumstances, precludes the appellate judge from participating in the review of the spouse's or relation's case.

Canon 3F. A remittal procedure provides the parties an opportunity to proceed without delay if they wish to waive the disqualification. To assure that consideration of the question of remittal is made independently of the judge, a judge must not solicit, seek, or hear comment on possible remittal or waiver of the disqualification unless the lawyers jointly propose remittal after consultation as provided in the rule. A party may act through counsel if counsel represents on the record that the party has been consulted and consents. As a practical matter, a judge may wish to have all parties and their lawyers sign the remittal agreement.

Technical Problem?
publicinformation@flcourts.org

the confidentiality of information received by a lawyer or judge participating in an approved lawyers assistance program; such an obligation, however, may be imposed by the rules of the program or other law.

Amended July 23, 1992, effective Jan. 1, 1993 (605 So.2d 252); effective Feb. 8, 2001 (795 So.2d 1); amended March 23, 2006, effective May 22, 2006 (SC04-2246), (933 So.2d 417). Amended April 12, 2012, effective July 1, 2012 (SC10-1967).

RULE 4-8.4 MISCONDUCT

A lawyer shall not:

(a) violate or attempt to violate the Rules of Professional Conduct, knowingly assist or induce another to do so, or do so through the acts of another;

(b) commit a criminal act that reflects adversely on the lawyer's honesty, trustworthiness, or fitness as a lawyer in other respects;

(c) engage in conduct involving dishonesty, fraud, deceit, or misrepresentation, except that it shall not be professional misconduct for a lawyer for a criminal law enforcement agency or regulatory agency to advise others about or to supervise another in an undercover investigation, unless prohibited by law or rule, and it shall not be professional misconduct for a lawyer employed in a capacity other than as a lawyer by a criminal law enforcement agency or regulatory agency to participate in an undercover investigation, unless prohibited by law or rule;

(d) engage in conduct in connection with the practice of law that is prejudicial to the administration of justice, including to knowingly, or through callous indifference, disparage, humiliate, or discriminate against litigants, jurors, witnesses, court personnel, or other lawyers on any basis, including, but not limited to, on account of race, ethnicity, gender, religion, national origin, disability, marital status, sexual orientation, age, socioeconomic status, employment, or physical characteristic;

(e) state or imply an ability to influence improperly a government agency or official or to achieve results by means that violate the Rules of Professional Conduct or other law;

(f) knowingly assist a judge or judicial officer in conduct that is a violation of applicable rules of judicial conduct or other law;

(g) fail to respond, in writing, to any official inquiry by bar counsel or a disciplinary agency, as defined elsewhere in these rules, when bar counsel or the agency is conducting an investigation into the lawyer's conduct. A written response shall be made:

(1) within 15 days of the date of the initial written investigative inquiry by bar counsel, grievance committee, or board of governors;

(2) within 10 days of the date of any follow-up written investigative inquiries by bar counsel, grievance committee, or board of governors;

(3) within the time stated in any subpoena issued under these Rules Regulating The Florida Bar (without additional time allowed for mailing);

(4) as provided in the Florida Rules of Civil Procedure or order of the referee in matters assigned to a referee; and

(5) as provided in the Florida Rules of Appellate Procedure or order of the Supreme Court of Florida for matters pending action by that court.

Except as stated otherwise herein or in the applicable rules, all times for response shall be calculated as provided elsewhere in these Rules Regulating The Florida Bar and may be extended or shortened by bar counsel or the disciplinary agency making the official inquiry upon good cause shown.

Failure to respond to an official inquiry with no good cause shown may be a matter of contempt and processed in accordance with rule 3-7.11(f) of these Rules Regulating The Florida Bar.

(h) willfully refuse, as determined by a court of competent jurisdiction, to timely pay a child support obligation; or

(i) engage in sexual conduct with a client or a representative of a client that exploits or adversely affects the interests of the client or the lawyer-client relationship.

If the sexual conduct commenced after the lawyer-client relationship was formed it shall be presumed that the sexual conduct exploits or adversely affects the interests of the client or the lawyer-client relationship. A lawyer may rebut this presumption by proving by a preponderance of the evidence that the sexual conduct did not exploit or adversely affect the interests of the client or the lawyer-client relationship.

The prohibition and presumption stated in this rule do not apply to a lawyer in the same firm as another lawyer representing the client if the lawyer involved in the sexual conduct does not personally provide legal services to the client and is screened from access to the file concerning the legal representation.

Comment

Lawyers are subject to discipline when they violate or attempt to violate the Rules of Professional Conduct, knowingly assist or induce another to do so, or do so through the acts of another, as when they request or instruct an agent to do so on the lawyer's behalf. Subdivision (a), however, does not prohibit a lawyer from advising a client concerning action the client is legally entitled to take, provided that the client is not used to indirectly violate the Rules of Professional Conduct.

Many kinds of illegal conduct reflect adversely on fitness to practice law, such as offenses involving fraud and the offense of willful failure to file an income tax return. However, some kinds of offense carry no such implication. Traditionally, the distinction was drawn in terms of offenses involving "moral turpitude." That concept can be construed to include offenses

concerning some matters of personal morality, such as adultery and comparable offenses, that have no specific connection to fitness for the practice of law. Although a lawyer is personally answerable to the entire criminal law, a lawyer should be professionally answerable only for offenses that indicate lack of those characteristics relevant to law practice. Offenses involving violence, dishonesty, or breach of trust or serious interference with the administration of justice are in that category. A pattern of repeated offenses, even ones of minor significance when considered separately, can indicate indifference to legal obligation.

A lawyer may refuse to comply with an obligation imposed by law upon a good faith belief that no valid obligation exists. The provisions of rule 4-1.2(d) concerning a good faith challenge to the validity, scope, meaning, or application of the law apply to challenges of legal regulation of the practice of law.

Subdivision (c) recognizes instances where lawyers in criminal law enforcement agencies or regulatory agencies advise others about or supervise others in undercover investigations, and provides an exception to allow the activity without the lawyer engaging in professional misconduct. The exception acknowledges current, acceptable practice of these agencies. Although the exception appears in this rule, it is also applicable to rules 4-4.1 and 4-4.3. However, nothing in the rule allows the lawyer to engage in such conduct if otherwise prohibited by law or rule.

Subdivision (d) of this rule proscribes conduct that is prejudicial to the administration of justice. Such proscription includes the prohibition against discriminatory conduct committed by a lawyer while performing duties in connection with the practice of law. The proscription extends to any characteristic or status that is not relevant to the proof of any legal or factual issue in dispute. Such conduct, when directed towards litigants, jurors, witnesses, court personnel, or other lawyers, whether based on race, ethnicity, gender, religion, national origin, disability, marital status, sexual orientation, age, socioeconomic status, employment, physical characteristic, or any other basis, subverts the administration of justice and undermines the public's confidence in our system of justice, as well as notions of equality. This subdivision does not prohibit a lawyer from representing a client as may be permitted by applicable law, such as, by way of example, representing a client accused of committing discriminatory conduct.

Lawyers holding public office assume legal responsibilities going beyond those of other citizens. A lawyer's abuse of public office can suggest an inability to fulfill the professional role of attorney. The same is true of abuse of positions of private trust such as trustee, executor, administrator, guardian, or agent and officer, director, or manager of a corporation or other organization.

A lawyer's obligation to respond to an inquiry by a disciplinary agency is stated in subdivision (g) of this rule and subdivision (h)(2) of rule 3-7.6. While response is mandatory, the lawyer may deny the charges or assert any available privilege or immunity or interpose any disability that prevents disclosure of a certain matter. A response containing a proper invocation thereof is sufficient under the Rules Regulating The Florida Bar. This obligation is necessary to ensure the proper and efficient operation of the disciplinary system.

Subdivision (h) of this rule was added to make consistent the treatment of attorneys who fail to pay child support with the treatment of other professionals who fail to pay child support, in accordance with the provisions of section 61.13015, Florida Statutes. That section provides for the suspension or denial of a professional license due to delinquent child support payments after all other available remedies for the collection of child support have been exhausted. Likewise, subdivision (h) of this rule should not be used as the primary means for collecting child support, but should be used only after all other available remedies for the collection of child support have been exhausted. Before a grievance may be filed or a grievance procedure initiated under this subdivision, the court that entered the child support order must first make a finding of willful refusal to pay. The child support obligation at issue under this rule includes both domestic (Florida) and out-of-state (URESA) child support obligations, as well as arrearages.

Subdivision (i) proscribes exploitation of the client or the lawyer-client relationship by means of commencement of sexual conduct. The lawyer-client relationship is grounded on mutual trust. A sexual relationship that exploits that trust compromises the lawyer-client relationship. Attorneys have a duty to exercise independent professional judgment on behalf of clients. Engaging in sexual relationships with clients has the capacity to impair the exercise of that judgment.

Sexual conduct between a lawyer and client violates this rule, regardless of when the sexual conduct began when compared to the commencement of the lawyer-client relationship, if the sexual conduct exploits the lawyer-client relationship, negatively affects the client's interest, creates a conflict of interest between the lawyer and client, or negatively affects the exercise of the lawyer's independent professional judgment in representing the client.

Subdivision (i) creates a presumption that sexual conduct between a lawyer and client exploits or adversely affects the interests of the client or the lawyer-client relationship if the sexual conduct is entered into after the lawyer-client relationship begins. A lawyer charged with a violation of this rule may rebut this presumption by a preponderance of the evidence that the sexual conduct did not exploit the lawyer-client relationship, negatively affect the client's interest, create a conflict of interest between the lawyer and client, or negatively affect the exercise of the lawyer's independent professional judgment in representing the client.

For purposes of this rule, a "representative of a client" is an agent of the client who supervises, directs, or regularly consults with the organization's lawyer concerning a client matter or has authority to obligate the organization with respect to the matter, or whose act or omission in connection with the matter may be imputed to the organization for purposes of civil or criminal liability.

Amended: July 23, 1992, effective Jan. 1, 1993 (605 So.2d 252); July 1, 1993 (621 So.2d 1032); July 1, 1993, effective. Jan. 1, 1994 (624 So.2d 720); Feb. 9, 1995 (649 So.2d 868); July 20, 1995 (658 So.2d 930); Sept. 24, 1998, effective Oct. 1, 1998 (718 So.2d 1179); Feb. 8, 2001 (795 So.2d 1); May 20, 2004 (SC03-705), 875 So.2d 448); December 8, 2005, the Supreme Court of Florida issued a revised version of its original October 6, 2005 opinion adopting this amendment, effective January 1, 2006 (SC05-206) (2005 WL 2456201), (916 So.2d 655); Amended March 23, 2006, effective May 22, 2006 (SC04-2246), (933 So.2d 417); Amended November 19, 2009, effective February 1, 2010 (SC08-1890) (34 Fla.L.Weekly S628a).

Rules for Certified and Court-Appointed Mediators

Contents

Part I Mediator Qualifications

Rule 10.100 Certification Requirements

(a) General. For certification as a county court, family, circuit court, dependency, or appellate mediator, a mediator must be at least 21 years of age and be of good moral character. For certification as a county court, family, circuit court or dependency mediator, one must have the required number of points for the type of certification sought as specifically required in rule 10.105.

(b) County Court Mediators. For initial certification as a mediator of county court matters, an applicant must have at least a high school diploma or a General Equivalency Diploma (GED) and 100 points, which shall include:

Committee Notes

2000 Revision. Mediated agreements will often impact persons or entities not participating in the process. Examples include lienholders, governmental agencies, shareholders, and related commercial entities. In family and dependency mediations, the interests of children, grandparents or other related persons are also often affected. A mediator is responsible for making the parties aware of the potential interests of such non-participating persons.

In raising awareness of the interests of non-participating persons, however, the mediator should still respect the rights of the parties to make their own decisions. Further, raising awareness of possible interests of related entities should not involve advocacy or judgments as to the merits of those interests. In family mediations, for example, a mediator should make the parents aware of the children's interests without interfering with self-determination or advocating a particular position.

Rule 10.330 Impartiality

(a) Generally. A mediator shall maintain impartiality throughout the mediation process. Impartiality means freedom from favoritism or bias in word, action, or appearance, and includes a commitment to assist all parties, as opposed to any one individual.

(b) Withdrawal for Partiality. A mediator shall withdraw from mediation if the mediator is no longer impartial.

(c) Gifts and Solicitation. A mediator shall neither give nor accept a gift, favor, loan, or other item of value in any mediation process. During the mediation process, a mediator shall not solicit or otherwise attempt to procure future professional services.

Committee Notes

2000 Revision. A mediator has an affirmative obligation to maintain impartiality throughout the entire mediation process. The duty to maintain impartiality arises immediately upon learning of a potential engagement for providing mediation services. A mediator shall not accept or continue any engagement for mediation services in which the ability to maintain impartiality is reasonably impaired or compromised. As soon as practical, a mediator shall make reasonable inquiry as to the identity of the parties or other circumstances which could compromise the mediator's impartiality.

During the mediation, a mediator shall maintain impartiality even while raising questions regarding the reality, fairness, equity, durability and feasibility of proposed options for settlement. In the event circumstances arise during a mediation that would reasonably be construed to impair or compromise a mediator's impartiality, the mediator is obligated to withdraw.

Subdivision (c) does not preclude a mediator from giving or accepting de minimis gifts or incidental items provided to facilitate the mediation.

Rule 10.340 Conflicts of Interest

(a) Generally. A mediator shall not mediate a matter that presents a clear or undisclosed conflict of interest. A conflict of interest arises when any relationship between the mediator and the mediation participants or the subject matter of the dispute compromises or appears to compromise the mediator's impartiality.

(b) Burden of Disclosure. The burden of disclosure of any potential conflict of interest rests on the mediator. Disclosure shall be made as soon as practical after the mediator becomes aware of the interest or relationship giving rise to the potential conflict of interest.

(c) Effect of Disclosure. After appropriate disclosure, the mediator may serve if all parties agree. However, if a conflict of interest clearly impairs a mediator's impartiality, the mediator shall withdraw regardless of the express agreement of the parties.

(d) Conflict During Mediation. A mediator shall not create a conflict of interest during the mediation. During a mediation, a mediator shall not provide any services that are not directly related to the mediation process.

(e) Senior Judge. If a mediator who is a senior judge has presided over a case involving any party, attorney, or law firm in the mediation, the mediator shall disclose such fact prior to mediation. A mediator shall not serve as a mediator in any case in which the mediator is currently presiding as a senior judge. Absent express consent of the parties, a mediator shall not serve as a senior judge over any case involving any party, attorney, or law firm that is utilizing or has utilized the judge as a mediator within the previous three years. A senior judge who provides mediation services shall not preside over the same type of case the judge mediates in the circuit where the mediation services are provided; however, a senior judge may preside over other types of cases (e.g., criminal, juvenile, family law, probate) in the same circuit and may preside over cases in circuits in which the judge does not provide mediation services.

Committee Notes

2000 Revision. Potential conflicts of interests which require disclosure include the fact of a mediator's membership on a related board of directors, full or part time service by the mediator as a representative, advocate, or consultant to a mediation participant, present stock or bond ownership by the mediator in a corporate mediation participant, or any other form of managerial, financial, or family interest by the mediator in any mediation participant involved in a mediation. A mediator who is a member of a law firm or other professional organization is obliged to disclose any past or present client relationship that firm or organization may have with any party involved in a mediation. The duty to disclose thus includes information relating

2017 WL 4422465

United States Court of Appeals,
Eleventh Circuit.

Raymond BERTHIAUME, Plaintiff–Appellant,
v.
David T. SMITH, individually, City of Key West, a Florida Municipal corporation, Defendants–Appellees.

No. 16–16345
|
(October 5, 2017)

Appeal from the United States District Court for the Southern District of Florida, D.C. Docket No. 4:15–cv–10001–JLK

Attorneys and Law Firms

Katherine E. Giddings, Ryan Daniel O'Connor, Akerman, LLP, Tallahassee, FL, Robert Dale Noll, Jr., Whitney Marie Untiedt, Akerman, LLP, Miami, FL, Hugh L. Koerner, Hugh L. Koerner, PA, Hollywood, FL, for Plaintiff–Appellant.

Michael Thomas Burke, Juan Marcos Martinez, Johnson Anselmo Murdoch Burke Piper & Hochman, PA, Fort Lauderdale, FL, for Defendants–Appellees.

Before HULL, JORDAN and GILMAN,* Circuit Judges.

Opinion

PER CURIAM:

*1 Raymond **Berthiaume** brought suit against Lieutenant David Smith of the **Key West** Police Department and the City of **Key West** ("the City") under 42 U.S.C. §§ 1983 and 1988 and Florida law, alleging claims of excessive force, false arrest, false imprisonment, battery/unnecessary force, and malicious prosecution, arising from Lieutenant Smith's October 2013 arrest of Berthiaume. Following a three-day trial, the jury returned a verdict in favor of the Defendants, and the district court subsequently denied Berthiaume's motion for a new trial.

On appeal, Berthiaume contends that he was denied a fair trial by an impartial jury. Berthiaume asserts, inter alia, that the district court abused its discretion in failing to ask jurors his proposed voir dire question, which was: "Do you harbor any biases or prejudices against persons who are gay or homosexual?" After review, and with the benefit of oral argument, we conclude that, given the particular facts and circumstances in this case described below, the district court abused its discretion in not asking that question. We explain why.

I. FACTUAL BACKGROUND

This case involved an altercation between two gay men who formerly had been partners. The following evidence was introduced at trial.

On the evening of October 26, 2013, Plaintiff **Berthiaume** attended the Fantasy Fest parade in **Key West**, Florida with his then-partner and now-husband, Jhon Villa, his friend Corey Smith, and his former partner, Nelson Jimenez. After the parade, **Berthiaume** and his companions remained in the area for the street party that followed. By the early morning hours of October 27, **Berthiaume**, Villa, and Smith were ready to go home and returned to their car, which was parked on a nearby street. Jimenez was not ready to leave and remained in one of the area gay bars.

After waiting for Jimenez by the car for some time, Berthiaume returned to the bar to find Jimenez and escort him back to the car so that the group could leave. As Berthiaume led Jimenez out of the bar with his hand on Jimenez's upper arm, Jimenez took the car keys from Berthiaume, twisted out of Berthiaume's grip, and ran down an adjacent alleyway. Berthiaume, clad only in boxer shorts or a loin cloth and flip flops,[1] followed Jimenez to retrieve the keys. During his pursuit of Jimenez, Berthiaume became frustrated and banged his hand against a street sign before continuing down the alleyway.

Lieutenant Smith and several other officers were on duty patrolling the Fantasy Fest area that night. Lieutenant

Smith and another officer observed the interaction between Berthiaume and Jimenez as they left the bar and believed that they were witnessing a fight or altercation between the two men. Lieutenant Smith testified that Berthiaume appeared to be swatting and grabbing at Jimenez with both hands as Jimenez tried to pull away, while the other officer testified that the only physical contact that occurred between the two men was Berthiaume's grasping of Jimenez's upper arm as he attempted to escort Jimenez back to the car. Although Berthiaume and one of his companions testified that Berthiaume was walking as he followed Jimenez down the alleyway, Lieutenant Smith and other officers testified that Berthiaume chased Jimenez down the alley and that both men were running.

***2** Lieutenant Smith and the other officers who were in the vicinity ran toward the alley to intervene. When he caught up to Berthiaume, Lieutenant Smith pushed Berthiaume in the shoulder to stop him from pursuing Jimenez, causing Berthiaume to fall to the ground. As a result of his fall, Berthiaume suffered a fractured wrist and jaw, both of which ultimately required surgery.

Lieutenant Smith spoke to Jimenez at the scene. Jimenez initially thanked Lieutenant Smith for intervening, but later stated that nothing wrong had happened, and he did not want to press charges against Berthiaume. According to Lieutenant Smith, Jimenez told him that he and Berthiaume were former partners and that they were trying to get back together.[2]

Despite Jimenez's unwillingness to press charges against Berthiaume, Lieutenant Smith chose to arrest Berthiaume and charge him with domestic battery.[3] Lieutenant Smith explained that "in domestic situations" such as this, "there is preferred arrest by the State of Florida" in order to ensure that the aggressor and victim are separated at least for the rest of the evening. Lieutenant Smith further indicated that an arrest was appropriate regardless of Jimenez's desire not to press charges because Lieutenant Smith personally had observed the battery on Jimenez. Lieutenant Smith also noted that victims of domestic battery sometimes "have different emotions for [the person who assaulted them] that make them not want to say something against that person because they don't want something bad to happen to them for their future."

II. MOTION FOR NEW TRIAL

The jury ultimately returned a verdict in favor of the defendants. Berthiaume filed a motion for new trial, arguing in part that he was deprived of a fair trial with an impartial jury when the district court refused to question the venire members regarding any potential bias they might have toward persons who are gay or homosexual. Berthiaume noted that homosexuals had only recently begun to gain acceptance in society, and many people still harbor bias or prejudice against homosexuals. Accordingly, Berthiaume contended that in a case such as his, involving both a gay party and gay witnesses, it is necessary for courts to inquire into prospective jurors' potential biases against homosexuals to ensure a fair trial.

The district court denied Berthiaume's motion for new trial.

III. DISCUSSION

In the context of racial bias, the Supreme Court has held that, under "special circumstances," the Constitution may require district courts to ask questions concerning racial bias—specifically, where racial issues are "inextricably bound up with the conduct of the trial," and there are substantial indications that racial prejudice would likely affect the jurors. Rosales–Lopez v. United States, 451 U.S. 182, 189–90, 101 S.Ct. 1629, 1635, 68 L.Ed.2d 22 (1981) (internal quotation marks omitted). The Supreme Court also indicated in Rosales–Lopez that questions regarding racial bias may still be warranted in the absence of such special circumstances, but that a district court's failure to ask such questions will be reversible only if the circumstances of the case indicate a reasonable possibility that racial prejudice might have influenced the jury. Id. at 190–91, 101 S.Ct. at 1635–36.

***3** In one unpublished case, this Court relied on this racial-bias precedent in holding that the district court reversibly erred by failing to inquire about sexual-orientation bias during voir dire. See United States v. Bates, 590 Fed.Appx. 882 (11th Cir. 2014). In Bates, the defendant was charged with numerous child pornography offenses. Id. at 883–84. In addition to finding child pornography on his computer, investigators found evidence indicating that Bates used the internet to meet other men for sexual encounters, as well as photographs of Bates engaged in sex acts with other men. Id. at 884–85. Bates sought to suppress that evidence, but the district court determined that it was relevant to establish his ownership and use of the computer. Id. Anticipating that the evidence of his private life would come out during trial, Bates then requested that the district court question prospective jurors "about any prejudice they might harbor against him on the basis of

his sexual activity with other men," which the district court declined to do. Id. at 884–85. During the trial, evidence regarding Bates's sexual activities with other men was "repeatedly paraded before the jury," and the jury ultimately convicted Bates on all counts. Id.

Here, as in Bates, Berthiaume's sexual orientation and that of his witnesses became "inextricably bound up with the issues to be resolved at trial." See id. at 887 (internal quotation marks and citation omitted). In describing the events leading up to Berthiaume's arrest, the witnesses repeatedly testified about Berthiaume's romantic relationships with Jimenez and Villa. Indeed, in explaining why he felt it necessary to arrest Berthiaume despite Jimenez's refusal to press charges, Lieutenant Smith explained that victims are often reluctant to press charges in "domestic situations" such as these because they have mixed emotions about the perpetrator.

Moreover, as in Bates, the district court here "did not ask any questions specific enough to determine whether any of the jurors might harbor prejudices against [Berthiaume] based on his sexual relationships." See Bates, 590 Fed.Appx. at 887. Nor were the district court's general inquiries regarding the jurors' ability to be impartial and its instruction that jurors not be prejudiced against witnesses based on the witnesses' backgrounds sufficient "to reach the important concerns highlighted by [Berthiaume's] proposed inquiry," as they were "broadly framed" and "not calculated to reveal latent prejudice." See id. (internal quotation marks omitted). As a result, the district court abused its discretion by failing to inquire about prejudice on the basis of sexual orientation during voir dire. Id.

Further, the Defendants have not shown "beyond a reasonable doubt that the error complained of did not contribute to the verdict obtained," and thus, Berthiaume is entitled to reversal. See id. at 888 (quoting Chapman v. California, 386 U.S. 18, 24, 87 S.Ct. 824, 828, 17 L.Ed.2d 705 (1967)). At the time of voir dire, the jury had no reason to know that Berthiaume's sexual orientation or that of his witnesses would be a part of the evidence at trial. Consequently, "they had no reason to offer up prejudices they might harbor on that basis when the District Court posed its general questions" regarding bias. See id. at 889. Moreover, the district court asked the jurors multiple questions about any biases or prejudices they might have against law enforcement. But the district court refused to ask any questions about prejudice on the basis of sexual orientation. Therefore, we have no way to discern whether the jury was biased against Berthiaume for that reason. See id. Given the repeated testimony at trial concerning Berthiaume's homosexual relationships with Villa and Jimenez, and the characterization of the altercation that led to Berthiaume's arrest as a domestic dispute, the risk that latent, undiscovered prejudices may have influenced the jury's verdict is substantial. Id.

IV. CONCLUSION

For the foregoing reasons, we vacate the district court's final judgment in favor of the Defendants and remand for a new trial.

VACATED AND REMANDED.[4]

All Citations

--- Fed.Appx. ----, 2017 WL 4422465

Footnotes

* Honorable Ronald Lee Gilman, United States Circuit Judge for the Sixth Circuit, sitting by designation.

1 The witnesses agreed that Berthiaume was shirtless and wearing flip flops at the time of his arrest, but differed in their descriptions of the rest of Berthiaume's attire. The police-officer witnesses described Berthiaume's outfit as a loin cloth, while Berthiaume and his companions stated that he was wearing boxer shorts.

2 Lieutenant Smith's arrest affidavit indicates that it was Berthiaume, rather than Jimenez, who provided this information to the officers.

3 The State subsequently declined to prosecute Berthiaume.

4 In light of our decision, we need not consider Berthiaume's alternative ground for a new trial based on the Defendant's exercise of peremptory challenges in alleged violation of Batson v. Kentucky, 476 U.S. 79, 106 S.Ct. 1712, 90 L.Ed.2d 69 (1986), and its

progeny.

End of Document

2017 WL 5616872
Only the Westlaw citation is currently available.
United States Court of Appeals,
Eleventh Circuit.

Raymond BERTHIAUME, Plaintiff–Appellant,
v.
David T. SMITH, Individually, and City of Key West, a Florida Municipal Corporation, Defendants–Appellees.

No. 16-16345
|
November 22, 2017

Synopsis

Background: Arrestee brought lawsuit under § 1983 and § 1988 against police officer, alleging claims including excessive force, false arrest, and malicious prosecution, arising out of his arrest for domestic battery. After jury returned verdict in favor of officer, arrestee filed motion for new trial, based on district court's refusal to question venire members regarding any potential bias they might have had against him on basis of his sexual orientation. The United States District Court for the Southern District of Florida, No. 4:15-cv-10001-JLK, denied arrestee's motion, and he appealed.

[Holding:] The Court of Appeals held that specific voire dire inquiry to determine whether any jurors might have harbored bias or prejudices against arrestee based on his sexual orientation was necessary to ensure that he was afforded fair trial with an impartial jury.

Vacated and remanded for new trial.

West Headnotes (4)

[1] **Jury**
Competence for Trial of Cause
Jury
Examination of Juror

The Constitution guarantees both criminal and civil litigants a right to an impartial jury, and voir dire can be an essential means of protecting this right. U.S. Const. Amend. 6.

Cases that cite this headnote

[2] **Constitutional Law**
Impartiality
Jury
Bias and prejudice

Although the conduct of voir dire is largely left to the sound discretion of the trial court, the district court's voir dire must at least provide reasonable assurance that prejudice will be discovered if present; this means that, in circumstances where juror prejudices are reasonably suspected on a particular subject, due process requires the court to ask questions on voir dire specifically to address that subject. U.S. Const. Amend. 14.

Cases that cite this headnote

[3] **Jury**
Bias and prejudice

To determine whether, during voire dire, specific questioning to disclose possible juror prejudices about a particular subject is necessary in a given case, courts look to whether, under all of the circumstances presented, there is a reasonable possibility that a particular type of prejudice might have influenced the jury.

Cases that cite this headnote

[4] **Jury**
Bias and prejudice

Specific voire dire inquiry to determine whether any of the jurors might have harbored bias or prejudices against arrestee based on his sexual

orientation was necessary to ensure that he was afforded a fair trial with an impartial jury in compliance with due process requirements in his civil suit against police officer alleging claims that included excessive force, false arrest, and malicious prosecution, arising out of his arrest for suspected domestic battery; given that it was obvious to the trial court and parties that case involved alleged domestic battery between former partners of the same sex, and that sexual orientation of arrestee and his witnesses would be central facts at trial that were inextricably bound up with issues to be resolved at trial, the risk that latent, undiscovered prejudices might have influenced jury's verdict for police officer was substantial. U.S. Const. Amends. 4, 6, 14.

Cases that cite this headnote

Appeal from the United States District Court for the Southern District of Florida, D.C. Docket No. 4:15-cv-10001-JLK

Attorneys and Law Firms

Katherine E. Giddings, Ryan Daniel O'Connor, Akerman, LLP, Tallahassee, FL, Hugh L. Koerner, Hugh L. Koerner, PA, Hollywood, FL, Robert Dale Noll, Jr., Whitney Marie Untiedt, Akerman, LLP, Miami, FL, for Plaintiff–Appellant.

Michael Thomas Burke, Juan Marcos Martinez, Johnson Anselmo Murdoch Burke Piper & Hochman, PA, Fort Lauderdale, FL, for Defendants–Appellees.

Ethan Rice, Hayley J. Gorenberg, Eric David Lesh, Lambda Legal Defense, New York, NY, for Amici Curiae Lambda Legal Defense and Education Fund, Inc., ACLU of Florida, National Association of Public Defense, Florida Association of Criminal Defense Lawyers, National LGBT Bar Association.

Before HULL, JORDAN and GILMAN,* Circuit Judges.

Opinion

PER CURIAM:

***1** Defendants David T. Smith and the City of Key West (collectively "defendants") moved for panel rehearing of an opinion originally filed on October 5, 2017 and reported at —— Fed.Appx. ——, 2017 WL 4422465. Defendants also moved for publication of the opinion. We grant in part and deny in part the defendants' motion for panel rehearing, grant the defendants' motion for publication of the opinion, vacate our prior opinion, and substitute for it the following opinion.

Raymond **Berthiaume** brought suit against Lieutenant David Smith of the **Key West** Police Department and the City of **Key West** ("the City") under 42 U.S.C. §§ 1983 and 1988 and Florida law, alleging claims of excessive force, false arrest, false imprisonment, battery/unnecessary force, and malicious prosecution, arising from Lieutenant Smith's October 2013 arrest of Berthiaume. Following a three-day trial, the jury returned a verdict in favor of the defendants, and the district court subsequently denied Berthiaume's motion for a new trial.

On appeal, Berthiaume contends that he was denied a fair trial by an impartial jury. Berthiaume asserts, inter alia, that the district court abused its discretion in failing to ask jurors his proposed voir dire question, which was: "Do you harbor any biases or prejudices against persons who are gay or homosexual?" After review, and with the benefit of oral argument, we conclude that, given the particular facts and circumstances in this case described below, the district court abused its discretion in not asking that question. We explain why.

I. FACTUAL BACKGROUND

This case involved an altercation between two gay men who formerly had been partners. The following evidence was introduced at trial.

On the evening of October 26, 2013, **Berthiaume** attended the Fantasy Fest parade in **Key West**, Florida with his then-partner and now-husband, Jhon Villa, his friend Corey Smith, and his former male partner, Nelson Jimenez. After the parade, **Berthiaume** and his companions remained in the area for the street party that followed. By the early morning hours of October 27, **Berthiaume**, Villa, and Smith were ready to go home and returned to their car, which was parked on a nearby street. Jimenez was not ready to leave and remained in one of the area gay bars.

After waiting for Jimenez by the car for some time, Berthiaume returned to the bar to find Jimenez and escort him back to the car so that the group could leave. As

Berthiaume led Jimenez out of the bar with his hand on Jimenez's upper arm, Jimenez took the car keys from Berthiaume, twisted out of Berthiaume's grip, and ran down an adjacent alleyway. Berthiaume, clad only in boxer shorts or a loin cloth and flip flops,[1] followed Jimenez to retrieve the keys. During his pursuit of Jimenez, Berthiaume became frustrated and banged his hand against a street sign before continuing down the alleyway.

***2** Lieutenant Smith and several other officers were on duty patrolling the Fantasy Fest area that night. Lieutenant Smith and another officer observed the interaction between Berthiaume and Jimenez as they left the bar and believed that they were witnessing a fight or altercation between the two men. Lieutenant Smith testified that Berthiaume appeared to be swatting and grabbing at Jimenez with both hands as Jimenez tried to pull away, while the other officer testified that the only physical contact that occurred between the two men was Berthiaume's grasping of Jimenez's upper arm as he attempted to escort Jimenez back to the car. Although Berthiaume and one of his companions testified that Berthiaume was walking as he followed Jimenez down the alleyway, Lieutenant Smith and other officers testified that Berthiaume chased Jimenez down the alley and that both men were running.

Lieutenant Smith and the other officers who were in the vicinity ran toward the alley to intervene. When he caught up to Berthiaume, Lieutenant Smith pushed Berthiaume in the shoulder to stop him from pursuing Jimenez, causing Berthiaume to fall to the ground. As a result of his fall, Berthiaume suffered a fractured wrist and jaw, both of which ultimately required surgery.

Lieutenant Smith spoke to Jimenez at the scene. Jimenez initially thanked Lieutenant Smith for intervening, but later stated that nothing wrong had happened, and he did not want to press charges against Berthiaume. According to Lieutenant Smith, Jimenez told him that he and Berthiaume were former partners and that they were trying to get back together.[2]

Despite Jimenez's unwillingness to press charges against Berthiaume, Lieutenant Smith chose to arrest Berthiaume and charge him with domestic battery.[3] Lieutenant Smith explained that "in domestic situations" such as this, "there is preferred arrest by the State of Florida" in order to ensure that the aggressor and victim are separated at least for the rest of the evening. Lieutenant Smith further indicated that an arrest was appropriate regardless of Jimenez's desire not to press charges because Lieutenant Smith personally had observed the battery on Jimenez. Lieutenant Smith also noted that victims of domestic battery sometimes "have different emotions for [the person who assaulted them] that make them not want to say something against that person because they don't want something bad to happen to them for their future."

II. MOTION FOR NEW TRIAL

The jury ultimately returned a verdict in favor of the defendants. Berthiaume filed a motion for new trial, arguing in part that he was deprived of a fair trial with an impartial jury when the district court refused to question the venire members regarding any potential bias they might have toward persons who are gay or homosexual. Berthiaume noted that homosexuals had only recently begun to gain acceptance in society, and many people still harbor bias or prejudice against homosexuals. Accordingly, Berthiaume contended that in a case such as his, involving both a gay party and gay witnesses, it is necessary for courts to inquire into prospective jurors' potential biases against homosexuals to ensure a fair trial.

The district court denied Berthiaume's motion for new trial.

III. DISCUSSION

[1] [2]"The Constitution guarantees both criminal and civil litigants a right to an impartial jury," and "voir dire can be an essential means of protecting this right." Warger v. Shauers, —— U.S. ——, 135 S.Ct. 521, 528–29, 190 L.Ed.2d 422 (2014). Although the conduct of voir dire is largely "left to the sound discretion of the trial court," the district court's voir dire must at least "provide reasonable assurance that prejudice will be discovered if present." United States v. Hill, 643 F.3d 807, 836 (11th Cir. 2011) (internal quotation marks omitted). This means that, in circumstances "where juror prejudices are reasonably suspected" on a particular subject, due process requires the court to ask questions on voir dire specifically to address that subject. See United States v. Ochoa–Vasquez, 428 F.3d 1015, 1037 (11th Cir. 2005); see also, e.g., Morgan v. Illinois, 504 U.S. 719, 735–36, 112 S.Ct. 2222, 2233, 119 L.Ed.2d 492 (1992) (holding that, in a capital case, trial courts must inquire about juror's views on the death penalty); Ham v. South Carolina, 409 U.S. 524, 529, 93 S.Ct. 848, 851, 35 L.Ed.2d 46 (1973) (holding that, under certain circumstances, trial courts must inquire about racial bias);

Jordan v. Lippman, 763 F.2d 1265, 1280–82 (11th Cir. 1985) (holding that, in a case involving considerable pre-trial publicity, the district court must specifically inquire about prospective jurors' exposure to the pre-trial publicity).

*3 [3]To determine whether specific questioning is necessary in a given case, courts look to whether, under all of the circumstances presented, there is a reasonable possibility that a particular type of prejudice might have influenced the jury. Rosales–Lopez v. United States, 451 U.S. 182, 190–92, 101 S.Ct. 1629, 1635–36, 68 L.Ed.2d 22 (1981); see also Ristaino v. Ross, 424 U.S. 589, 596–98, 96 S.Ct. 1017, 1021–22, 47 L.Ed.2d 258 (1976) (concluding that specific questioning about racial bias was not necessary where the particular circumstances of the case "did not suggest a significant likelihood that racial prejudice might infect [the] trial"). "The critical factor" in making this determination is whether the potentially prejudicial issue is "inextricably bound up with the conduct of the trial," such that there is a "consequent need, under all the circumstances, specifically to inquire into [the] possible [specific] prejudice in order to assure an impartial jury." Rosales–Lopez, 451 U.S. at 189, 101 S.Ct. at 1635 (internal quotation marks omitted).

[4]Here, at the outset, it was obvious to the district court and the parties that this case involved an alleged domestic battery between former partners of the same sex, and that the sexual orientation of Berthiaume and his witnesses would be central facts at trial and were "inextricably bound up" with the issues to be resolved at trial. Id. at 189, 101 S.Ct. at 1635–36. First, the complaint refers to Villa as Berthiaume's "partner" and to Jimenez as his "ex-partner." The complaint alleges that Berthiaume and his partner Villa wanted to leave a public area and retire for the evening, but ex-partner Jimenez wanted to stay, took Villa's keys and ran. Berthiaume ran after Jimenez to get the keys. The complaint further states that in the arrest affidavit, Lieutenant Smith falsely stated that the dispute between Berthiaume and Jimenez arose because they were "trying to get back together," a fact Berthiaume denied. Moreover, the parties stipulated in their joint pretrial statement, and the arrest affidavit confirmed, that Lieutenant Smith charged Berthiaume with "domestic battery," which Berthiaume disputed. Berthiaume's lawsuit challenged Lieutenant Smith's false basis for arresting him and Lieutenant Smith's use of force to effectuate that arrest. So the relationship between Berthiaume and Jimenez, and the nature of the alleged domestic dispute between them as former partners, was of critical importance.

Further, after the district court ruled against Berthiaume's request for a supplemental voir dire question on sexual orientation bias, Berthiaume's counsel raised a Batson[4] challenge to the defendants' use of peremptory challenges against two potential jurors who he perceived as gay, explaining to the district court that "the evidence in this case will show that my client and all of his witnesses are gay." Based on the district court's sidebar with the attorneys concerning the Batson challenge, it was clear at that point that additional questioning on sexual orientation bias was necessary.

But the district court here did not ask any questions to determine whether any of the jurors might harbor prejudices against Berthiaume based on his sexual orientation. Nor were the district court's general inquiries regarding the jurors' ability to be impartial and its instruction that jurors not be prejudiced against witnesses based on the witnesses' backgrounds sufficient to reach the important concerns highlighted by Berthiaume's proposed inquiry because the general inquiries were broadly framed and not calculated to reveal latent prejudice. See Morgan, 504 U.S. at 734–36, 112 S.Ct. at 2232–33 (explaining that, where a reasonable possibility of prejudice on a specific subject exists, general questions about fairness and impartiality are insufficient to address the specific concern). As a result, under the particular facts and circumstances of this case, the district court abused its discretion by failing to inquire about prejudice on the basis of sexual orientation during voir dire. See id.

*4 Further, under these facts and circumstances, there is a "reasonable possibility that [sexual orientation bias] might have influenced the jury." See Rosales–Lopez, 451 U.S. at 192, 101 S.Ct. at 1636. Given the long history of cultural disapprobation and prior legal condemnation of same-sex relationships, the risk that jurors might harbor latent prejudices on the basis of sexual orientation is not trivial. See Obergefell v. Hodges, ––U.S. ––––, 135 S.Ct. 2584, 2596–97, 192 L.Ed.2d 609 (2015). Just two terms ago in Obergefell, the Supreme Court noted that for much of the 20th century, homosexuality was considered a mental illness, and same-sex intimacy was prohibited by law in many states. Id. at 2596. And despite the more recent "shift in public attitudes toward greater tolerance," Obergefell itself is evidence that issues regarding homosexuality continue to be debated in our society. See id. at 2597. While some jurors are not biased based on sexual orientation, some realistically are.

In contrast to the parties and the court, the jury, at the time of voir dire, had no reason to know that this case involved an alleged domestic dispute between former partners of the same sex or that Berthiaume's sexual

orientation or that of his witnesses would be a part of the evidence at trial. Consequently, they had no reason to offer up prejudices they might harbor on that basis when the district court posed its general questions regarding bias. Moreover, the district court here asked the jurors multiple questions about any biases or prejudices they might have against law enforcement. But the district court refused to ask any questions at all about prejudice on the basis of sexual orientation. Therefore, we have no way to discern whether the jury was biased against Berthiaume for that reason. Given the pretrial documentation concerning Berthiaume's homosexual relationships with both Villa and Jimenez, and the characterization of the altercation that led to Berthiaume's arrest as a domestic dispute, the risk that latent, undiscovered prejudices may have influenced the jury's verdict is substantial.

And this is not a case where that pretrial notice that sexual orientation was a pivotal fact proved false; instead, that notice proved well-founded. Indeed at trial, in describing the events leading up to Berthiaume's arrest, the witnesses repeatedly testified about Berthiaume's romantic relationships with Jimenez and Villa and what role, if any, the relationships played in the dispute between Berthiaume and Jimenez on the night in question. Additionally, in explaining why he felt it necessary to arrest Berthiaume despite Jimenez's refusal to press charges, Lieutenant Smith explained that victims are often reluctant to press charges in "domestic situations" such as these because they have mixed emotions about the perpetrator. Thus, the district court's error in this case was not harmless, and Berthiaume is entitled to reversal. See Rosales–Lopez, 451 U.S. at 192, 101 S.Ct. at 1636.

As explained above, whether specific voir dire questioning is required in a given case is a fact-specific inquiry which looks to the totality of the circumstances presented in that case and whether the district court had notice of the nature of the dispute. See Rosales–Lopez, 451 U.S. at 189–92, 101 S.Ct. at 1635–36. Thus, although we conclude that the district court erred in refusing to inquire about sexual orientation bias in this particular case, we do not hold that refusal to do so would be an abuse of discretion in every case. Rather, whether a district court abuses its discretion by failing to ask a proposed voir dire question about sexual orientation bias will depend on the subject matter and issues in the case and what notice the district court has that issues of sexual orientation will potentially be a central part of the evidence at trial. And if the court has notice, parties requesting such questioning during voir dire should aid the district court by explaining, at the time the request is made, the factual basis for the requested question.

IV. CONCLUSION

***5** For the foregoing reasons, we vacate the district court's final judgment in favor of the defendants and remand for a new trial.

VACATED AND REMANDED.[5]

All Citations

--- F.3d ----, 2017 WL 5616872

Footnotes

* Honorable Ronald Lee Gilman, United States Circuit Judge for the Sixth Circuit, sitting by designation.

1 The witnesses agreed that Berthiaume was shirtless and wearing flip flops at the time of his arrest, but differed in their descriptions of the rest of Berthiaume's attire. The police-officer witnesses described Berthiaume's outfit as a loin cloth, while Berthiaume and his companions stated that he was wearing boxer shorts.

2 Lieutenant Smith's arrest affidavit indicates that it was Berthiaume, rather than Jimenez, who provided this information to the officers.

3 The State subsequently declined to prosecute Berthiaume.

4 Batson v. Kentucky, 476 U.S. 79, 106 S.Ct. 1712, 90 L.Ed.2d 69 (1986).

5 In light of our decision, we need not consider Berthiaume's alternative ground for a new trial based on the defendants' exercise

of peremptory challenges in alleged violation of Batson and its progeny.

End of Document

2017 WL 4018836

NOTICE: THIS OPINION HAS NOT BEEN RELEASED FOR PUBLICATION IN THE PERMANENT LAW REPORTS. UNTIL RELEASED, IT IS SUBJECT TO REVISION OR WITHDRAWAL.

District Court of Appeal of Florida,
First District.

Shawnest Angelo IVEY, Appellant,
v.
STATE of Florida, Appellee.

CASE NO. 1D15–5803
|
Opinion filed September 13, 2017.

Synopsis

Background: Defendant was convicted in the Circuit Court, Jackson County, Shonna Young Gay, J., of possession of XLR11 and methamphetamine. Defendant appealed.

Holdings: The District Court of Appeal, Makar, J., held that:

[1] counsel was not required to object to factual basis of peremptory challenge, and

[2] counsel preserved objection to peremptory challenge.

Reversed and remanded.

Winsor, J., filed opinion concurring in part and dissenting in part.

West Headnotes (2)

[1] **Criminal Law**
Competency of jurors and challenges

Defense counsel was not required to object to factual basis of prosecutor's race-neutral justification for peremptory challenge to sole African American on jury panel, that potential juror allegedly gave prosecutor a "dirty look" after overhearing prosecutor's joke regarding potential juror in hallway, to preserve issue for review, where neither trial judge nor defense counsel were present when potential juror allegedly gave the "dirty look," there was no other evidence to establish "the look," and prosecutor failed to raise "the look" until late in the jury selection process.

Cases that cite this headnote

[2] **Criminal Law**
Competency of jurors and challenges

Defense counsel preserved objection to prosecutor's race-neutral justification for peremptory challenge that potential juror allegedly gave prosecutor "dirty look" after overhearing prosecutor's joke regarding potential juror in hallway, where defense counsel timely raised an objection to peremptory challenge and confirmed objection just prior to jury being sworn in response to judge's specific reference to peremptory challenge regarding potential juror.

Cases that cite this headnote

An appeal from the Circuit Court for Jackson County. Shonna Young Gay, Judge.

Attorneys and Law Firms

Andy Thomas, Public Defender, Jennifer P. LaVia, Special Public Defender, for Appellant.

Pamela Jo Bondi, Attorney General, Michael McDermott, Assistant Attorney General, for Appellee.

Opinion

MAKAR, J.

*1 Shawnest Angelo Ivey, who was convicted of

possession of XLR11 (commonly referred to as K2, Spice, or Synthetic Marijuana) and methamphetamine, argues that the trial court erred in allowing a peremptory challenge to strike an African–American juror where the State's race-neutral reason—that the juror gave the prosecutor a dirty look—was neither observed by the trial court nor supported by the record. We agree and reverse.[1]

I. Background

Juror number 46 is an African–American woman. During jury selection, neither the State nor defense counsel questioned her other than to ask whether she was employed (she said yes). After the close of voir dire, but before finalizing the jury, the prosecutor informed the trial court that the State wanted to use a peremptory strike on the potential juror. In response, defense counsel requested a race-neutral reason for the strike because juror number 46 was the only African–American on the jury panel.

According to the prosecutor, the basis for the peremptory strike was that she had made a joke about the potential juror during a break in jury selection between a prior case and Ivey's case. She said the potential juror overheard her and then gave "a look" that the prosecutor believed reflected bias against her. Neither the trial judge, defense counsel, nor anyone else observed the described encounter; nor was there any record evidence to establish "the look" other than the prosecutor's statement. The trial court accepted the State's proffered reason for the strike as race-neutral and removed juror number 46 from further consideration.

Immediately thereafter, the trial court asked the State and defense counsel if they were agreeable to the jury members who had been selected. Defense counsel said he went over the entire panel with Ivey, who said he "agrees and accepts this jury." Just prior to swearing the jury, however, the trial court and defense counsel had the following exchange:

> Court: The only additional thing is looking at my seating chart for jury selection yesterday, I had seated for juror number 46, and just for record purposes, wanted to make sure she was not a cause, she was a peremptory challenge. And there was a challenge race neutral reason given, and she was excused based on the state using a peremptory challenge. With that, is there anything else we need to address this morning before we bring the jury in?
>
> Defense: Your Honor, the only thing other than—everything you said is fine. What I would like to do, I've made a few objection in preliminary proceedings and objected to evidence and objected to different things. I would like to just make that as a continuing objection, so they don't come back and say we failed to object in the trial.
>
> Court: I will just state for record purposed [sic], any ruling that has already been made by me, I recognized [defense counsel's] continue [sic] objections, that has been the ruling that has been made by the Court.

***2** The jury was sworn and Ivey was later found guilty on the possession charges. On appeal, he argues that the trial court erred in allowing the State to use its peremptory strike on the potential juror.

II. Analysis

At the outset, the State does not argue that the trial judge's substantive ruling on the State's peremptory strike was correct, probably because it is directly contrary to Dorsey v. State, 868 So.2d 1192 (Fla. 2003), which held that a proponent of a peremptory strike "based on nonverbal behavior may satisfy its burden of production of a race-neutral reason ... only if the behavior is observed by the trial court or otherwise has record support." Id. at 1199. Here, "the look" was neither observed by the trial judge nor does it have record support (other than the prosecutor's say-so, which Dorsey says is inadequate). As such, the State makes two preservation arguments: (a) Ivey failed to dispute the factual basis of the race-neutral reason offered by the State, and (b) Ivey failed to renew his objection prior to the jury being sworn.

A. *Dorsey*–Footnote 3

[1]First, the State argues that the challenge to its peremptory strike is unpreserved because defense counsel did not challenge the factual basis for the State's proffered race-neutral reason, citing a footnote in Dorsey that the genuineness of a race-neutral reason offered by a party "does not arise where the opponent does not dispute the observation proffered as the reason for the strike." 868 So.2d at 1196 n.3. The footnoted language, however, relates to situations—such as in Dorsey—in which the nonverbal behavior occurs where the judge and defense counsel are physically present and able to observe it,

which was not the case here.

The issue in Dorsey was "whether a party's observation of a juror's nonverbal behavior may constitute a genuine, race-neutral reason for a peremptory challenge when the purported behavior is challenged by the opposing party and was neither observed by the trial court nor otherwise supported by the record." Id. at 1194. In Dorsey, the prosecutor believed that a prospective juror showed a lack of interest during attorney questioning and moved to strike her on that basis; defense counsel disagreed and countered that the prospective juror was the only one who showed interest and enthusiasm for being on the jury. Id. The trial judge, who was present throughout juror questioning, did not observe the alleged indifference, but took the prosecutor "at her word" in claiming it occurred, thereby upholding the strike. Id. at 1194–95. The supreme court overturned that ruling, holding that the record must have support—other than the prosecutor's own perception of the nonverbal conduct—to meet the burden imposed under Melbourne v. State, 679 So.2d 759 (Fla. 1996), and its progeny. Dorsey, 868 So.2d at 1200.

The supreme court in Dorsey rejected the view that the testimony of the prosecutor was itself sufficient, even when the judge and others are present, saying, "if the proponent of a strike were permitted to meet its burden of production based solely on an attorney's subjective, uncorroborated, and disputed impression of a juror's demeanor, the appellate court would have no basis to determine if the trial court's decision to accept the explanation was clearly erroneous." Id. at 1200. Ivey's counsel made a similar point, saying, "[t]he only thing we have is [the prosecutor] telling us this at sidebar what happened and I have no way to show that this—I just don't think it's a race neutral reason in the record and that's what she has to have."

***3** As is evident, the situation in Dorsey involved a prosecutor's interpretation of events that occurred during voir dire, where defense counsel was present and had an observable basis for disagreeing with the prosecutor's assertion. That is why the supreme court in Dorsey cited both Watson v. State, 841 So.2d 659 (Fla. 4th DCA 2003) and Carter v. State, 762 So.2d 1024 (Fla. 3d DCA 2000), each involving juror behavior that was observed in the courtroom during the jury selection process. Under those circumstances, if defense counsel is present, observes the juror's nonverbal conduct, and "does not dispute the observation proffered as the reason for the strike," the admonition in footnote three would apply. 868 So.2d at 1196 n.3.

Unlike Dorsey, however, neither the trial judge nor the defense counsel in Ivey's case were present when juror number 46 allegedly gave "the look" that the prosecutor says occurred. For this reason, Ivey's counsel could not have disputed the prosecutor's unilateral claim that the juror gave her the stink-eye during a break in the proceedings. Indeed, Ivey's counsel pointed out that the prosecutor failed to raise "the look" until late in the jury selection process, making it doubly difficult to assess the matter ("if [the prosecutor] thought that that was a problem, she should have brought it up during voir dire where we could have addressed it").

Because defense counsel did not have the benefit of observing the juror's nonverbal behavior, it was impossible for him to object to the factual basis for the prosecutor's assertion. As such, defense counsel was not required to disagree with the prosecutor's subjective interpretation of events under Dorsey.

B.

[2]The State also argues that Ivey failed to preserve the issue for review because his "continuing objection" prior to the jury being sworn was inadequate, relying on Mobley v. State, 100 So.3d 1170 (Fla. 1st DCA 2012). Ivey counters that the "continuing objection" was made directly in response to the trial judge confirming for record purposes the various objections raised, including the strike of juror number 46, which the trial court specifically raised and discussed.

This case differs from Mobley because the race-neutral reason offered in Mobley was based, at least in part, on interactions with the juror that occurred during voir dire, and was therefore on the record with defense counsel present. The State in Mobley offered a valid race-neutral reason for its strike, such that it was entitled to the presumption that its peremptory challenge was used in a nondiscriminatory manner; defense counsel was then required to "voic[e] a specific objection that the reason [was] not genuine or [was] not a pretext." Id. at 1172. The defendant in Mobley did neither—thereby failing to meet his burden to rebut the presumption to which the State was entitled for having offered a valid race-neutral reason. In this case, however, the State is not entitled to that presumption because the reason it offered was facially invalid under Dorsey: the nonverbal conduct of juror number 46 was neither witnessed by the trial court nor had record support, apart from the prosecutor's unilateral belief. See Dorsey 868 So.2d at 1199.

This case is more similar to Denis v. State, 137 So.3d 583

(Fla. 4th DCA 2014), where defense counsel objected to the State's use of a peremptory challenge, and the State's race-neutral justification offered for the strike was that the juror was "dozing off while the judge was speaking." Id. at 584. Defense counsel didn't observe the juror dozing and asked the trial judge whether he had witnessed it; the judge said he did not, but that that he would "take [the State]'s word for it" and upheld the strike. Id. In reversing, the Fourth District noted that "[p]reservation of an objection to the use of a peremptory strike requires more than one objection: an objection to initiate a *Melbourne* inquiry and another objection before the jury is sworn in." Id. at 585. As to the initiating objection, it held that the defendant had sufficiently preserved the issue for review by contesting the factual basis for the State's race-neutral reason because he "inform[ed] the court that defense counsel did not see the juror fall asleep and ask[ed] the court to address whether it witnessed this behavior." Id. at 585. This, the Fourth District stated, was enough to put the trial court "on notice that defense counsel questioned the genuineness of the state's reason for the challenge." Id. As to confirming the objection before the jury was sworn, the appellate court noted that "[i]n addition to making a proper objection to initiate a *Melbourne* inquiry, defense counsel preserved her objection to the state's peremptory challenge before the jury was sworn by accepting the jury 'subject to prior objections.' " Id.

***4** Here, Ivey's counsel timely raised an objection to the State's improper peremptory challenge and sufficiently confirmed that objection just prior to the jury being sworn in response to the trial judge's specific reference to the "juror number 46" issue. No doubt exists that the trial court knew of and was apprised of the specifics of Ivey's objection prior to the swearing of the jury, which was far different from the situation in Joiner v. State, 618 So.2d 174, 176 (Fla. 1993), where defense counsel "affirmatively accepted the jury immediately prior to its being sworn without reservation of his earlier-made objection." No affirmative acceptance was made immediatley prior to the jury being sworn that amounted to Ivey's counsel jettisoning his prior objections; instead, he renewed them in direct response to the trial court's inquiry. Because the State's stated reason was insufficient on its face under Dorsey, and Ivey adequately preserved the issue, we reverse and remand for a new trial.

REVERSE and REMAND.

BROWN, JOHN, ASSOCIATE JUDGE, concurs, and WINSOR, J. concurs in part and dissents in part with opinion.

WINSOR, J., concurring in part and dissenting in part.

I agree that we should affirm as to the trial court's order denying Ivey's motion to suppress. I would also affirm as to the disputed peremptory challenge because Ivey failed to preserve it.

After raising the objection initially, Ivey's counsel affirmatively told the court that Ivey "agrees and accepts this jury." Counsel's acceptance of the jury would lead "to a reasonable assumption that he had abandoned, for whatever reason, his earlier objection." *Joiner v. State*, 618 So.2d 174, 176 (Fla. 1993). The only question is whether counsel's discussion with the court the following day was sufficient to re-raise the issue. In the exchange, which the majority accurately recites in full, counsel said he'd "made a few objection (sic) in preliminary proceedings and objected to evidence and objected to different things." He then said he "would like to just make that as a continuing objection." That was not enough.

Reiterating objections to unspecified "different things" does not preserve a specific objection to a peremptory challenge. Up to the point of his general objection, Ivey had argued an unsuccessful motion in limine, an unsuccessful motion to dismiss, and an unsuccessful motion to suppress (which included an evidentiary hearing with contested evidentiary rulings). If Ivey sought to renew an objection on the peremptory issue, he needed more specificity. *Cf. Harrell v. State*, 894 So.2d 935, 940 (Fla. 2005) (noting that statutory requirement that objections be "sufficiently precise" to "fairly apprise[ ] the trial court of the relief sought and the grounds therefor," § 924.051(1)(b), Fla. Stat., is consistent with Florida Supreme Court holdings on preservation). Because Ivey did not specifically renew his abandoned objection before the jury was sworn, his claim is not preserved. *See Zack v. State*, 911 So.2d 1190, 1204 (Fla. 2005). I would therefore affirm.

All Citations

--- So.3d ----, 2017 WL 4018836, 42 Fla. L. Weekly D2004

Footnotes

1 As to Ivey's claim that the trial court erred in denying his motion to suppress his statements to police, we affirm.

STRIKE THREE:
DISCRIMINATION, INCENTIVES AND EVALUATION

Christopher A. Parsons, Johan Sulaeman, Michael C. Yates and Daniel S. Hamermesh*

This Version: August 2009

*Assistant professor of finance, University of North Carolina at Chapel Hill; assistant professor of finance, Southern Methodist University; assistant professor of finance, Auburn University; Sue Killam professor of economics, University of Texas at Austin, professor of labor economics, Maastricht University, and research associate IZA and NBER. We thank Jason Abrevaya, Jeff Borland, J.C. Bradbury, Jay Hartzell, Larry Kahn, William Mayew, Gerald Oettinger, Allen Sanderson, Alan Schwarz, John Siegfried, Paul Tetlock, Sheridan Titman, Deborah White and Justin Wolfers, participants at seminars at several universities, and the editor and four referees for helpful discussions.

Abstract

Major League Baseball umpires express their racial/ethnic preferences when they evaluate pitchers. Strikes are called less often if the umpire and pitcher do not match race/ethnicity, but mainly where there is little scrutiny of umpires. Pitchers understand the incentives and throw pitches that allow umpires less subjective judgment (e.g., fastballs over home plate) when they anticipate bias. These direct and indirect effects bias performance measures of minorities downward. The results suggest how discrimination alters discriminated groups' behavior generally. They imply that biases in measured productivity must be accounted for in generating measures of wage discrimination.

I. Introduction and Motivation

Tests of labor market discrimination typically compare labor market outcomes (e.g., wages, promotion rates) across groups and, after controlling for worker productivity, assign any residual differences to discrimination. But what if an evaluator who discriminates along the dimension being studied *subjectively* determines a worker's measured productivity, as is true in all but the simplest piece-rate environments? A worker subjected to such biased evaluations might appear less productive, which ordinarily would justify a lower wage. However, in this case the econometrician would underestimate, or perhaps even miss altogether, instances of labor market discrimination when they in fact exist.

A subtler complication is that workers, anticipating biased evaluations, may alter their behavior in ways intended to minimize its impact. For example, consider a police officer who can either: 1) write traffic citations (the number of which can be objectively measured), or 2) investigate crimes (which is subject to performance review by a higher-ranking officer). If the officer has sufficient discretion, a biased evaluation in the second activity would lead the officer to alter the allocation of her time. Presumably, a positive bias would cause the officer to spend more time investigating crimes, and vice versa. Such bias-induced shifts in behavior further complicate the identification problem in assessing the impact of discrimination in labor markets.

This study addresses both of these issues, using detailed data on the evaluation, observed strategies and performances of Major League Baseball (MLB) players. Our focus is on racial/ethnic bias, specifically between the umpire (evaluator) and the pitcher (worker), although the arguments we develop apply to any type of subjective bias.[1] We pay particular attention to

[1]Pitches are only subject to the umpire's discretion (are "called") when the batter does not swing, rendering necessary a judgment of whether the pitch was a "ball" or a "strike."

the race/ethnicity "match" of the umpire and pitcher, which occurs when, for example, a Black umpire evaluates a Black pitcher, as opposed to evaluating a White or Hispanic pitcher.

Our first observation is that pitchers who match the race/ethnicity of the home-plate umpire appear to receive slightly favorable treatment, as indicated by a higher probability that a pitch is called a strike, compared to players who do not match. Although this confers an advantage to some players at the expense of others, the effect we document here is small, on average affecting less than a pitch per game. Much more interesting are situations *when* and *where* the effects are strongest. Roughly one-third of the ballparks we study contain a system of computerized cameras (QuesTec) used to evaluate the umpires, comparing their ball/strike calls to a less subjective standard. Umpires have strong incentives to suppress any bias in such situations, as the QuesTec evaluations are important for their own career outcomes. With such *explicit* monitoring, evidence of any race or ethnicity preference vanishes entirely.

We find similar effects with *implicit* monitoring; when a game is well attended (and presumably more closely scrutinized), or when the pitch is pivotal for an at-bat, race/ethnicity matching again plays no role in the umpire's evaluation. In situations where the umpire is neither explicitly nor implicitly monitored, the effect of the bias is considerable. As an example, a Hispanic pitcher facing a Hispanic umpire in a low-scrutiny setting (e.g., no cameras, poorly attended) receives strikes on 32.5 percent of called pitches, which drops to 30.0 percent if a Black umpire is behind the plate.

However, such direct effects are magnified when pitchers adjust their strategies in response to biased evaluations. Like the multi-tasking police officer mentioned above, a pitcher can alter his behavior to make himself either more immune, or more exposed to, the umpire's judgment. Specifically, pitches thrown near the borders of the strike zone (e.g., over one of

home plate's corners) are called balls nearly as frequently as they are called strikes. They constitute a "fuzzy" region where the umpire can employ maximum subjectivity. Because such pitches are more difficult for batters to hit than those thrown directly into the strike zone, we would expect pitchers aware of favorable treatment to throw disproportionately to this fuzzy region. We find exactly this. Pitchers who match the umpire's race/ethnicity attempt to "paint the corners," throwing pitches allowing umpires the most discretion. This tendency is much stronger in low-scrutiny situations, when umpires face a lower cost of indulging their preferences.

At the end of both exercises, we are left with two specific conclusions. First, incentives matter. Unless provided strong incentives not to do so, umpires appear to allow the pitcher's race or ethnicity to influence their subjective judgments. This leads to a small, but non-trivial, direct effect on the game, simply by increasing the probability that a pitch is called a strike. Second, pitchers appear to understand these incentive effects, and take measures to protect themselves by avoiding situations requiring high subjectivity when facing a downward bias.

The results also lead to two general conclusions. First, these results show that when worker productivity is measured subjectively, and when such measurements are biased by discrimination, the usual tests for discrimination are biased toward finding nothing. We illustrate the size of this bias in our sample of baseball pitchers. Second, they illustrate the need to be aware of the manner in which discrimination in one facet of evaluation can lead market participants to alter their behavior in other dimensions.

Baseball offers several advantages when studying discrimination. First, because every pitch is potentially subject to the home-plate umpire's discretion when it is thrown (several hundred times per game), there is sufficient scope for racial/ethnic discrimination to be expressed

as well as for it to affect games' outcomes significantly. In addition, the very large number of independent pitch-level observations involving the interaction of different races/ethnicities allows us not only to explore umpires' preference for players of their own race/ethnicity, but also to examine preferences toward other races/ethnicities.[2] An additional feature of baseball data is that, unlike other sports where a group dynamic among officials may alter the expression of individual biases, the home-plate umpire is exclusively responsible for calling every pitch in a typical baseball game.[3]

The most fortunate aspect of the data set is that it allows us to develop several independent proxies for the scrutiny of the umpire's decisions, and in so doing, to test for the existence of price-sensitive discrimination by umpires. The time period that we analyze, 2004-2008, is special, because only during this time were a portion of the ballparks outfitted with computers and cameras to monitor umpires' balls and strikes calls. Because umpires are randomly assigned to venues, observing differences in their behavior between parks with and without monitoring technology makes a convincing case that properly placed incentives can have the desired effect. These results allow us not only to describe how biases can influence subjective performance valuations, but also to offer prescriptive suggestions to minimize their impact.

Several studies (e.g., Luis Garicano *et al*, 2005; Eric W. Zitzewitz, 2006, and Thomas J. Dohmen, 2008) have examined home-team preferences by referees/judges in various sports, and another, Michael A. Stoll *et al* (2004) examines racial match preferences in employment

[2]The data also include a small number of Asian pitchers, but because there are no Asian umpires, we exclude them in our analysis. Given their trivial numbers however, their inclusion gives nearly identical results in every instance.

[3]Umpires can be positioned behind home plate or at first, second or third base. The home-plate umpire occasionally appeals to either the first- or third-base umpire, but this is a relatively infrequent occurrence, and in any case is usually initiated by the home-plate umpire himself to help determine if the batter swung at the ball.

generally. Our study most closely resembles Joseph Price and Justin J. Wolfers' (2007) work on NBA officiating crews' racial preferences. Although the first part of our empirical analysis corroborates their findings (but for a different sport), we are mainly interested in *when* or *where* racial/ethnic bias is most likely to be observed. Here, we offer two insights. First, we show that discrimination is price-sensitive, so that making it more costly reduces its expression. Second, we show that, when quantifying how players are affected by biased performance evaluations, the direct effect is only part of the story. Because players will alter their strategies in response, even situations that are seemingly insulated from a biased evaluator (e.g., non-called pitches in baseball games) are affected.

This research adds to a large literature on racial discrimination in sports, specifically in baseball, going back at least to Anthony H. Pascal and Leonard A. Rapping (1972), James D. Gwartney and Charles T. Haworth (1974) and Gerald W. Scully (1974), and recently J.C. Bradbury (2007) generally, with others dealing with particular racial/ethnic issues (Clark Nardinelli and Curtis J. Simon, 1990, David W. Findlay and Clifford E. Reid, 1997, and Rodney D. Fort and Andrew M. Gill, 2000). It includes studies of such outcomes as productivity, wages, customers' approbation of players, selection for honors, and others. There is some evidence of wage disparities among baseball players of different races, but the results are mixed, e.g., Lawrence M. Kahn (1991). The conclusions of racial discrimination (or lack thereof) in this literature depend upon each player's productivity being accurately measured, as measured productivity is typically the crucial control variable. We suggest questioning this central assumption: If officials' judgments are themselves subject to racial/ethnic bias, adjusting for differences in the returns to *measured* productivity will not enable us to obtain proper measures of the extent of discrimination.

The results allow us to think about the deeper question of measuring discrimination generally. If, as we show here, the match to the race/ethnicity of their evaluator affects evaluations of workers, then the measured productivity of the worker will depend on the nature of that match. This difficulty has serious implications for measuring discrimination and is another manifestation of the difficulty of identifying discrimination pointed out by Stephen G. Donald and Daniel S. Hamermesh (2006).

In the next section we describe the pitch- and game-level data and explain our classification of umpires' and players' races/ethnicities. We analyze individual pitches in Section III, and in Section IV we show that umpires express these preferences strongly only in times of low scrutiny. We examine the indirect impact of discrimination on pitchers' strategies in Section V. Section VI shows the overall effects on pitchers' performances and derives the size of the effects of biased performance evaluation on the measurement of wage discrimination generally and for the example of pitchers' salaries.

II. Data

*Pitches***.** There are 30 teams in Major League Baseball, with each team playing 162 games in each regular season. During a typical game each team's pitchers throw about 150 pitches, so that approximately 700,000 pitches are thrown each season. We collect pitch-by-pitch data from ESPN.com for every regular-season MLB game from 2004-2008.[4] Our final dataset consists of 3,524,624 total pitches. For each pitch we identify the pitcher, pitcher's team, batter,

[4]The pitch-by-pitch information is from: http://sports.espn.go.com/mlb/playbyplay?gameId=NNNNNNNN&full=1, where NNNNNNNN represents the nine-digit game ID. The first six digits correspond to the year, month and date of the game. The box score information is from http://sports.espn.go.com/mlb/boxscore?gameId=NNNNNNNN .

batter's team, catcher, pitch count, score, inning, and pitch outcome. We classify each pitch into one of seven exhaustive and mutually exclusive categories: Called strike, called ball, swinging strike, foul, hit into play, intentional ball or hit by pitch. We supplement each pitch observation with other relevant information, including the stadium name, home team, away team, and the identities and positions of all four umpires.

Player and Umpire Race/Ethnicity. We next classify each position player, pitcher and umpire who appears in our dataset as White, Hispanic, Black or Asian. To begin this task, we collect country of birth for every player and umpire. Players or umpires are classified as Hispanic if they were born in: Colombia, Cuba, Curacao, Dominican Republic, Mexico, Nicaragua, Panama, Puerto Rico or Venezuela. Players from Japan, South Korea and Taiwan are classified as Asian. We classify an additional 69 players using an *AOL Sports* article which lists every African-American player on a MLB roster at the beginning of the 2007 season.[5] We also utilize a similar list of past and present Hispanic players in MLB from Answers.com.[6] All remaining unclassified players and umpires are classified by visual inspection of pictures found in Internet searches.[7] Three of the race/ethnic groups are represented among umpires (there are no Asian umpires in MLB), and all four are represented among pitchers.

Table 1 presents the distributions of the pitch outcomes. The first row of the table summarizes all pitches, while subsequent rows sub-divide pitches based on the race/ethnicity of

[5]The complete list can be found at http://Blackvoices.aol.com/Black_sports/special/_a/african-american-players-in-mlb/20070413095009990001.

[6]The complete list can be found at http://www.answers.com/topic/list-of-hispanic-players-in-major-league-baseball.

[7]For a few umpires, no pictures were available on the internet. For each of them we watched past games in which the umpires worked to ascertain their race/ethnicity. Any such classification is necessarily ambiguous in a number of cases. To the extent that we have inadvertently classified pitchers, umpires, or batters in ways different from how they might be treated on the field, this will introduce measurement error into the matches and thus reduce the strength of any results that we generate.

the pitcher, the batter and the home plate umpire, respectively. Approximately 46 percent of pitches elicit a swing from the batter, hit the batter, or are intentionally thrown out of the strike zone. Our pitch-level analysis focuses on the 54 percent of pitches (1.89 million) that result in called strikes or balls, since these alone are subject to evaluation by the home-plate umpire. Of these, about 32 percent are called strikes, and the rest are called balls.

The table also reports the number of pitchers, batters and home-plate umpires in each of the four race/ethnicity categories. The percentages of White pitchers (70 percent) and batters (61 percent) are lower in our sample than the percentage of White umpires (89 percent). On the other hand, Hispanics, comprising 23 percent of pitchers and 26 percent of batters, are under-represented among umpires (only 5 percent). Black pitchers, batters and umpires comprise 3 percent, 10 percent, and 6 percent of the samples, respectively. Asian players comprise 3 percent of pitchers and 2 percent of batters.

Pitch Location. For approximately one-third of the games played in 2007 season and all of those played in 2008, we collected from PITCHf/x several additional variables. PITCHf/x, a computerized technology owned by Sportvision, uses two cameras to record the path of a pitch from the pitcher's hand to home plate.[8] The parameters measured and calculated using this technology include: 1) the pitch type, determined using MLB's proprietary neural net classification algorithm, 2) the estimated pitch location when it crosses the home plate relative to the center of the front of the home plate, and 3) the top and bottom of the strike zone as determined by the PITCHf/x operator.[9]

[8]The information is provided by the MLB free of charge at: http://gd2.mlb.com/components/game/mlb/.

[9]The operator sets a horizontal line at each batter's belt as he settles into the hitting position, and the PITCHf/x software adds four inches up to define the top of the zone. For the bottom of the zone, the PITCHf/x operator sets a horizontal line at the hollow of each batter's knee. More information on PITCHf/x's parameters can be found at: http://fastballs.wordpress.com/category/pitchfx-glossary/ and http://webusers.npl.illinois.edu/~a-nathan/pob/tracking.htm.

Pitcher Performance. For each starting pitcher's appearance in each game, we collect from box scores the number of innings pitched, the numbers of hits, runs and home runs allowed, walks, strikeouts, and earned runs (downloaded from the ESPN website). We also obtain the final score of the game to identify the winning and losing teams.

III. Called Pitches and Umpire-Pitcher Matches

Table 2 reports for each pitcher/umpire racial/ethnic combination the number of pitches thrown, the number of called pitches, the number of called strikes and the percentage of called pitches that are strikes. About two-thirds of the called pitches in our sample occur when the umpire and pitcher share the same race/ethnicity (mostly White pitcher/White home-plate umpire). While the percentage of pitches that are called is similar in situations where the umpire's and pitcher's race/ethnicity match and in situations where they do not (53.7 percent), a central difference is that the percentage of called pitches that are strikes is higher when they match (32.0 percent) than when they do not (31.5 percent).

The summary statistics in Table 2 ignore possible differences inherent in the quality or "style" of pitchers by race/ethnicity. They also ignore the possibly different outcomes generated by non-random assignment of pitchers to face different opponents, and of umpires to games played by particular teams.[10] To account for these and other potential difficulties, our central test for umpires' discrimination in calling strikes is the specification:

$$I(\text{Strike} \mid \text{Called Pitch})_i = \gamma_0 + \gamma_1 UPM_i + \gamma_2 Controls_i + \varepsilon_i, \quad (1)$$

[10]Examination of umpires' schedules indicates that, while umpires typically travel as a four-person crew throughout much of the year, crews are randomly assigned across teams, ballparks, geography, and league (American or National). Furthermore, umpires rotate in a specific order, i.e., each serves as the home-plate umpire exactly every fourth game, resulting in random assignment of umpires to starting pitchers.

where the dependent variable is an indicator of whether a called pitch is a strike, the γ are parameters, ε is a well-behaved error term, and *i* indexes pitches. The main explanatory variable of interest is UPM, an indicator of whether the umpire (U) and pitcher (P) match (M) on race/ethnicity. In almost all of our tests, we include fixed effects for each pitcher, umpire and batter so that UPM picks up the *marginal* effect of a racial/ethnic match between the home-plate umpire and pitcher. That is, because any player or race-specific effects are swept out by the fixed effects, umpires' bias is identified purely via the interaction term, UPM.

In addition to these, we employ a number of control variables. Pitch-count indicators, which record how many balls and strikes have accrued during a particular at-bat, are crucial because pitchers alter the location of their pitches based on the ball-strike count. Inning indicators are also included, because pitchers are usually less fatigued early in games, and because a "relief" pitcher often replaces a pitcher who starts the game in later innings, with a different (often reduced) accuracy.[11] Home-field bias is captured by top-of-the-inning indicators**,** which account for which team is pitching. Lastly, we include the pitcher's score advantage (defined as the number of runs, potentially negative, by which the pitcher's team is ahead).

Table 3 presents the results of estimating equations where the pitcher's and umpire's race/ethnicity are allowed to influence the likelihood of a called strike. All the estimates are based on linear-probability models (but probit estimates present the same picture) with heteroskedasticity-robust standard errors. The first three columns show specifications separately for White, Black and Hispanic pitchers, respectively, controlling for umpire race/ethnicity and pitcher fixed effects. The next three columns show separate equations for White, Black and Hispanic umpires, respectively, controlling for pitcher race/ethnicity and umpire fixed effects.

[11]With pitcher fixed effects, this second reason for inning indicators is obviously subsumed.

The final three columns include all pitchers and umpires, with each column adding successive vectors of fixed effects, including in the final column pitchers, umpires and batters.

There is some, albeit weak, evidence of favoritism by umpires for pitchers who match their race/ethnicity. For example, Column (1) shows that Hispanic umpires judge White pitchers more harshly than do White umpires (the omitted indicator variable), but that they judge Hispanic pitchers more favorably (Column (3)). Similarly, Column (4) shows that White umpires, the overwhelming majority, judge minority pitchers more harshly than they judge White pitchers. Taking the results in Column (9) with the full sets of control variables and fixed effects as the best description of the underlying behavior, however, it is quite clear that there is no generally significant impact of the match on umpire evaluations (p=.34).

Although the results with the broadest sets of fixed effects do not suggest a significant effect of the umpire-pitcher match, the point estimate implies that a given called pitch is approximately 0.16 percentage points more likely to be a strike if the umpire and pitcher match race/ethnicity. The likelihood that a given called pitch is called a strike is 31.9 percent. Thus when the umpire matches the pitcher's race/ethnicity, the rate of called strikes rises by one-half percent above the rate when there is no match.[12]

IV. Biased Evaluation When Bias Is Costly

One might examine the results in Table 3 and conclude that, while the point estimates are interesting, their statistical insignificance means that there is very little here. Given an economist's view that agents acting out their preferences will react to the price of an activity, however, it is worthwhile examining the impacts of umpire-pitcher matches as the price of

[12]As a check on this issue we re-estimated the model including sequentially the race/ethnic match between the first-, second- and third-base umpire and the pitcher. None of these extensions materially changes our conclusions.

discrimination changes. We begin by asking what factors affect the price of expressing racial or ethnic discrimination. Studies of cognitive behavior indicate that presenting the biased party with counter-examples of the stereotype of interest can reduce the severity and/or frequency of the biased behavior (Stephanie A. Goodwin *et al*, 2000; Irene V. Blair, 2002). In other words, simply making conscious a sub-conscious bias imposes a sufficient psychological cost to mitigate its expression. Another mechanism is to increase the visibility of the biased party's behavior, potentially exposing the offender to social or legal penalties. Here we proxy the price of discrimination by the extent to which an umpire's evaluations of pitchers will be scrutinized. We employ three different measures to examine whether a higher price of discrimination reduces the extent to which umpires engage in discriminatory behavior.

The first source of scrutiny is QuesTec, a computerized monitoring system intended to evaluate the accuracy and consistency of home-plate umpires' judgments. From 2004-2008, MLB had installed QuesTec in 11 of its 30 ballparks.[13] QuesTec's Umpire Information System (UIS) consists of four cameras that track and record the location of each pitch, providing information about the accuracy and precision of each umpire's ball and strike calls. Despite opposition from some umpires and players (perhaps most memorably, pitcher Curt Schilling's assault on a camera after a poor outing), the QuesTec system served as an important tool to evaluate umpires during our sample period. According to the umpires' union's agreement with MLB, QuesTec is the primary mechanism to gauge umpire performance. If more than 10 percent of an umpire's calls differ from QuesTec's records, his performance is considered

[13]The ballparks of the Anaheim Angels, Arizona Diamondbacks, Boston Red Sox, Cleveland Indians, Oakland Athletics, Milwaukee Brewers, Houston Astros, New York Mets, Tampa Bay Devil Rays, Chicago White Sox, and New York Yankees.

substandard, which can influence his promotion to "crew chief," assignment to post-season games, or even retention in MLB.[14]

Because QuesTec is installed in roughly one-third of ballparks, and because umpiring crews are rotated randomly around the league's ballparks, virtually every umpire in our data set calls a substantial number of pitches in parks with and without QuesTec.[15] Additionally, both the umpires' and teams' schedules change every year, exposing each umpire to a wide cross-section of batters and pitchers in both types of parks. Throughout the analysis we test whether greater scrutiny—the possibly higher cost of bias in subjective evaluation of pitches in QuesTec parks—leads umpires to call strikes "by the book." Any role that racial/ethnic (or any other) preferences play in influencing pitch calls should be mitigated if costs of being judged substandard are imposed, as through QuesTec. Some pitchers may, however, react differently from others in response to QuesTec.[16] For that reason, in all of the estimates in this part (and hereafter) we include fixed effects not only for each pitcher, umpire and batter, but also for the presence or absence of QuesTec in each game, i.e., pitcher-QuesTec fixed effects, umpire-QuesTec fixed effects, and batter-umpire-QuesTec fixed effects.

Figure 1 graphs the average percentages of called pitches that are strikes in ballparks with and without QuesTec, for White and minority pitchers respectively. The effect of monitoring on umpires' behavior is apparent, with both White and minority pitchers being judged differently by

[14]An umpire's evaluation is not based solely on QuesTec. If an umpire falls below the QuesTec standards, his performance is then reviewed by videotape and live observation by other umpires to determine his final evaluation score. No such measures are taken, however, if an umpire meets the QuesTec standards.

[15]The fraction of games in which QuesTec was installed was virtually identical for all umpires in our sample, differing for the few umpires calling only a handful of games.

[16]For example, New York Mets pitcher Tom Glavine, known as a "finesse" pitcher who depends on pitches close to the strike zone border, complained publicly that QuesTec's influence on umpire calls forced him to change his style (Associated Press, July 9, 2003). Glavine reports that he was told, "[umpires do] not call pitches on the corners at Shea [his home ballpark] because they [the umpires] don't want the machine to give them poor grades."

umpires of the matched race/ethnicity, depending on whether the pitch is thrown in a park with QuesTec installed. The difference in the called-strike percentage between QuesTec and non-QuesTec parks is significant for both White and minority pitchers.

Table 4 contains the results of estimating (1) separately for QuesTec and non-QuesTec parks, with controls for inning, pitch count, pitcher score advantage, and top of the inning.[17] The results are striking: In ballparks with the UIS, shown in Column (1), the coefficient on UPM is -0.48 percentage points and is not significantly different from zero. In parks without QuesTec, shown in Column (2), the same coefficient is 0.59 percentage points per pitch (p=.007). These differences make clear why UPM is not significant in the aggregate sample. The effects found in Table 3 averaged the statistically significant positive impact of an unscrutinized match (non-QuesTec) with a statistically insignificant negative impact of a scrutinized match (QuesTec) that is nearly as large. Thus, in the presence of price-sensitive discrimination, we should expect the point estimates in Table 3 to be low, since the entire sample consists of a mix of high- and low-scrutiny games. Specifically, QuesTec covers about 37 percent of pitches, so that the average result from Table 3 is easily reconciled: (.37)(-.48)+(.63)(.59)=.19, close to the 0.16 estimate obtained with a comparable set of fixed effects.

Column (3) of Table 4 presents the results when the QuesTec indicator is interacted with UPM. When the pitcher and umpire match race/ethnicity, pitching in a QuesTec park reduces the likelihood that a called pitch is ruled a strike by over 1 percentage point, more than offsetting the favoritism shown by umpires when QuesTec does not monitor them. Each effect is highly significant, implying that umpires implicitly allow their apparent preference for matched pitchers to be expressed when the pitches underlying their decisions are not recorded.

[17]The direct effect of being in a QuesTec park is, of course, not directly observable, being subsumed in the pitcher-QuesTec fixed effects.

QuesTec is an explicit monitoring technology. Implicit monitoring can have similar effects, suggesting that even subtle incentive mechanisms can have desirable effects on otherwise discriminatory outcomes. The two measures for implicit scrutiny of umpires are crowd attendance (scaled by stadium capacity) and the "importance" of the pitch.[18]

The idea for the first is simple. Having many fans close to home plate presumably exposes the umpire to their scrutiny—a badly called pitch is unlikely to go unnoticed.[19] Figure 2 confirms that crowd attendance, like QuesTec, dramatically alters umpire behavior. A game is defined as "well-attended" if the crowd attendance is above the median percentage capacity in this sample, roughly 70 percent. Compared to well-attended games, umpires calling poorly-attended games appear to favor pitchers of matched race/ethnicity. In the case of White pitchers, both minority and White umpires tend to call fewer strikes in poorly-attended games, but the reduction in strikes called by minority umpires is over three times larger. The same effect is seen to an even greater degree among minority pitchers. During well-attended games, matching minority umpires call about 0.8 percent *fewer* strikes. They call 0.7 percent *more* strikes in poorly attended ones, a net effect over 1.5 percentage points.

In Columns (1) and (2) of Panel A in Table 5, we show the results of estimating (1) separately for well- and poorly-attended games respectively. Each equation includes the same battery of controls as in Table 4, i.e., pitcher, umpire and batter fixed effects, pitch counts, and inning indicators. As with the QuesTec results, the UPM variable is significant (p=.008) only in

[18]We scale by stadium capacity to minimize the impact of differences between stadium sizes. If we assume that stadiums populate relatively uniformly, attendance/capacity is a good proxy for the *number* of fans close enough to judge pitch location. In any case, this scaling makes little difference in our results. If instead we use attendance, all coefficients of interest remain highly significant.

[19]Percentage attendance may also proxy the popularity of the participating teams or the importance of a particular game. Thus, not only might the umpire be exposed to more scrutiny from the additional fans present at well-attended games, but he may also face added scrutiny in the form of larger television audiences and increased air-time given to game highlights.

poorly-attended games, with an effect of 0.64 percentage points per pitch. During well-attended games there is no significant effect of an umpire-pitcher racial/ethnic match and, as before, the point estimate is negative. Column (3) generalizes the results by aggregating all games and interacting UPM with the indicator for a game being well attended. Compared to a pitch in a poorly-attended game when the umpire and pitcher do not match, a pitch called by an umpire of the same race/ethnicity as the pitcher is 0.36 percentage points more likely to be judged a strike. If the game is well-attended, a pitch is no more likely to be called a strike if the pitcher and umpire match race/ethnicity. The results for this completely different proxy for the price of discrimination are qualitatively identical to those obtained for the QuesTec/non-QuesTec distinction.

A third proxy for the scrutiny of umpires varies many times within each game. We separate pitches into two categories, "terminal" and "non-terminal." A pitch is potentially terminal if the umpire's next judgment can terminate the batter's plate appearance. Specifically, a pitch that is thrown with two strikes and/or three balls is potentially terminal, as a third strike or fourth ball terminates the at-bat. In such situations, the umpire's judgment is likely to be scrutinized more heavily by the pitcher, batter, catcher, managers and fans. An initial glimpse into the effects of this distinction is shown in Figure 3. Here we observe the same contrast as for the previous two proxies for scrutiny, as umpires appear to favor pitchers with whom they match only in non-terminal counts, when scrutiny is likely to be reduced.

Columns (4) and (5) of Panel B of Table 5 show estimates of (1) separately for terminal and non-terminal pitches, with pitcher, umpire and batter fixed effects and the usual set of control variables. We consider pitches of differing importance separately, with the result that the coefficients of UPM have opposite signs. For pitches that cannot be terminal, the estimated

coefficient of UPM is 0.31 percentage points (p=.15)—umpires favor pitchers who match their own race/ethnicity. For potentially terminal pitches, where scrutiny of the umpire is likely to be greater, umpires appear to judge pitchers of their own race/ethnicity (insignificantly) more harshly than unmatched pitchers. In Column (6) all pitches are aggregated, and UPM is interacted with an indicator for potentially terminal pitches. The results mimic those implicit in the estimates in Columns (4) and (5), as the coefficient on the interaction term is negative and significant at better than the 1 percent level.

In Columns (7) and (8) we consider another source of within-game variation in implicit scrutiny. We assume that, because umpires' evaluations are more likely to be pivotal late in games, scrutiny in the first few innings is likely to be comparatively less. We thus designate the first third (three innings) of a game as "early," and the remainder "late." We expect that a terminal count will have a stronger effect on the outcome of a pitcher-umpire racial/ethnic match in early innings. Comparing the results across the two columns, we see that this is the case, with the magnitude of the interaction between terminal count and UPM being over twice as large in early as in late innings (-0.86 vs. -0.38 percentage points).

Our proxies for scrutiny are not redundant. The correlation between QuesTec and attendance percentage is small, and because the type of pitch (terminal or non-terminal) is a within-game measure, it is necessarily uncorrelated with either between-game measure. It is therefore not surprising that, when all three interactions are included simultaneously in Panel C, everything remains significant with nearly identical magnitudes as in Panels A and B.

Before proceeding to issues of robustness, we briefly address whether the UPM effect is due to positive bias for pitchers who match the umpire's race/ethnicity (i.e., favoritism), or to negative bias against those who do not match. Answering this question in our context is

difficult, because ball and strike calls are inherently subjective. (Compare this to tennis, where the definition of a shot being "in" or "out" is completely objective, allowing, for example, computerized instant replay to reverse the judge's calls.) Absent an objective standard on strike calls, we cannot precisely quantify the bias' direction; but comparing umpires' behavior between QuesTec and non-QuesTec ballparks provides some illumination.

If one accepts the premise that umpires exercise special care in QuesTec parks, the strike percentage there, although not perfect, is closer to the desired benchmark of objectivity that would permit the desired calculation. For each of the nine possible race/ethnic combinations, we compare the called strike percentage in QuesTec parks (the quasi-objective benchmark) to that in non-QuesTec parks. First, all three cases of a match (e.g., White-White) show a higher called strike percentage in non-QuesTec parks, which suggests favoritism in less scrutinized situations. Second, five of the six cases of non-match show a lower called strike percentage in non-QuesTec parks, which suggests negative bias. Such a two-sided pattern not only justifies the use of an aggregate UPM variable in Tables 3-5; it also demonstrates that the effect is symmetric and pervasive across nearly every possible combination. However, we do not focus further on the positive/ negative bias distinction, because baseball – and indeed all games with winners and losers – is a zero-sum game. It is *relative* treatment that matters most, just as in labor markets generally it is disparate treatment, not the difficult-to-identify distinction between the absence of favoritism and the presence of negative bias, that underlies so much case law.

Other matches. An umpire influenced by the race of the pitcher may also be influenced by that of the batter or the catcher, especially because in the latter case, the umpire is in continuing close contact. We find little evidence to support this argument. In the same types of regressions as in Tables 3-5, but with new matching variables, there is some very weak evidence

that batters receive the type of preferential treatment experienced by pitchers. But the magnitudes and statistical significance are much smaller in every case. Reliable evidence for similar catcher-umpire dynamics is even weaker. Umpires appear focused on the pitchers they are judging. Their matches with other relevant players do not affect their judgments.

Post-season. The three scrutiny proxies we employ have the advantage of splitting the sample of called pitches into two large groups, generating the statistical power required to detect subtle differences in called strike probabilities. There are many additional cross-sectional tests one could perform, e.g., comparing playoff to regular season games (because the former are likely to be particularly scrutinized), but such thin cross-sectional comparisons contain almost no power. For example, we replicate the analysis in Panel C of Table 5, aggregating playoff and regular season games, and including interaction terms for post-season pitches with the coefficients of interest (unreported). There is only the weakest of evidence that playoff situations reduce further the expression of umpire bias (the interaction of post-season with UPM is negative, as expected, but the p-value is 0.74). We encounter a similar problem when, for example, examining particularly "important" games, such as those pivotal for playoff races late in the season.

Umpire and City Characteristics. It may be that umpires' measurable characteristics (beyond their race/ethnicity) and those of the city where a game is played explain our results. We collected demographic information on each umpire from a variety of sources and include his age and experience, and in many cases both his state of birth and residence. For each ballpark we also obtain the racial/ethnic breakdown of the surrounding metropolitan statistical area.

We find no evidence that the racial composition of an umpire's birthplace or residence predicts his propensity to penalize non-matching players, but there is some weak evidence that

bias is more likely among younger and less experienced umpires. The coefficient on UPM in a re-specification of (1) among the upper half of umpires ranked by experience is less than half its magnitude in estimates for umpires in the lower half of the distribution. If (1) is re-estimated separately for the 18 "crew chiefs," veterans selected for their seniority and performance, the point estimate of the coefficient on UPM is nearly zero. This evidence is consistent with models of selection or learning. Perhaps discriminating umpires are not promoted and are dropped from the ranks. Alternatively, experience may teach umpires to restrain their own biases.

We also re-estimated the basic equation for Blacks, and for Hispanics, separately, adding in each case main effects and interactions with UPM of the percentage of the minority group in the metropolitan area where the ballpark is located. Among Blacks the interaction was positive, but statistically insignificant; among Hispanics it was negative, and also statistically insignificant. Our conclusions are not affected by the racial/ethnic mix of the team's catchment area. [20]

Gaming the System. Perhaps managers are implicitly both aware of these preferences and able to act upon them. Because the majority of umpires are White, there is a distinct advantage for a team with one or more minority pitchers (particularly starting pitchers) to have QuesTec in its home park. We found no information about the how teams were awarded QuesTec in their home parks, or whether they could influence this choice. A second possibility is that teams

[20]The overwhelming majority of minority pitchers are Hispanic. We have aggregated them, but some are White Hispanics, while others are Black Hispanics. To allow for the possibility that the two different groups of minority umpires might treat Hispanic pitchers who match their own characteristics differently from other Hispanic pitchers, we visually inspected the pitchers' pictures, divided the Hispanic aggregate into White and Black groups, and consequently re-defined UPM. This reclassification had almost no effect on the estimates produced in Tables 3-5. Implicitly, Hispanic and other umpires treat Hispanic pitchers the same regardless of the pitcher's racial identity. We also investigated whether American-born Hispanic pitchers were treated differently from Hispanic pitchers born outside the U.S., and found no evidence that the pitcher's birthplace affected expressed racial/ethnic bias by umpires.

receiving QuesTec systems traded for minority pitchers from teams whose parks were not similarly equipped.

Although we have no direct evidence, some simple calculations suggest that either possibility may have merit. For visiting pitchers, the percentage of pitches thrown in QuesTec parks is nearly identical for Whites and minorities (37.4 and 37.9 respectively). This is to be expected, because on average, teams play approximately the same fraction of opponents whose home stadiums contain QuesTec. Thus, there is no evidence that visiting managers adjust their pitching lineups to minimize the exposure of their minority pitchers to the subjective bias of a White umpire.

Home pitchers tell a different story. Minority home pitchers throw 39.2 percent of their pitches in QuesTec parks, compared to only 35.5 percent for White pitchers. Home minority pitchers are more likely to be in QuesTec environments, which can only be the case if their home ballpark has QuesTec. This is consistent with either initial non-random assignment of QuesTec to teams with a disproportionate number of minority pitchers, with transactions that increase the fraction of minority pitchers for teams already equipped with QuesTec, or with game-time lineup juggling by home teams. Although we cannot distinguish among these alternatives, this evidence is interesting by suggesting that biased evaluations in one area (e.g., called strikes) may have unintended consequences in other arenas (e.g., the allocation of Minority pitching talent). Note that none of these possibilities alters the significance or interpretation of the previous results, as all regressions control for player ability, umpire tendencies, and the presence or absence of QuesTec.

V. The Effects of Biased Evaluations on Agents' Strategies

The pitch-level evidence makes very clear that direct effects on pitch outcomes are small. Of course, one can construct specific examples where the estimated direct effect is fairly large: a Black pitcher throwing a *non-terminal* pitch in the *early innings* of *poorly-attended* games in a *non-QuesTec* ballpark gains over 6 percentage points by matching (41.4 vs. 35.2 percent called strikes). But in most situations, the direct impact on called pitches is not large.

Indirect effects on players' strategies may, however, have larger impacts on the outcomes of plate appearances and games. The dynamic between a pitcher and batter is clearly affected by each party's beliefs about the umpire's evaluation in the event of a called pitch. If a pitcher expects favoritism, he will incorporate this advantage into his strategy, perhaps throwing pitches that allow the umpire more discretion. This in turn may change the batter's optimal behavior. If the batter expects such pitches to be called strikes, he is forced to swing at "worse" pitches, which reduces the likelihood of getting a hit.[21]

To appreciate more fully such induced changes in strategy, for all the starting pitchers for whom such data are available (over 500,000 pitches), we augment the pitch-level data with the data set on pitch characteristics.[22] This level of detail allows addressing the extent to which pitchers alter their strategies (e.g., location and type of pitch), when facing a biased subjective evaluation. Panel A of Table 6 summarizes the two location variables of interest: 1) the horizontal pitch distance, and 2) the pitch height. The first is the distance (in feet) from the

[21]Examining the coefficients on the count indicators in Table 3 illustrates the intuition. When the pitcher has a substantial advantage in the count, he has little incentive to throw a "hittable" pitch, i.e., one near the middle of the plate. Instead, he usually throws pitches near the corners that are both less likely to be hit if the batter swings, and less likely to be called strikes if the batter does not. Such behavior translates into sizeable advantages for pitchers depending on the count. In 2004, batters got a hit 33 percent of the time when the count was 2-1 (two balls and one strike), but dropped to less than 18 percent when the count was 1-2.

[22]The number of pitches differs slightly across the panels because of difficulties in classifying by location and type. Jowei Chen (2007) used these data for a single season as controls to examine racial bias in MLB umpires' calls.

center of home plate. (The slightly negative mean value for this variable reflects the tendency to avoid hitting or pitching inside to batters, and most batters are right-handed.) The second is calculated as the pitch's vertical distance from the center of the strike zone, which is set by the computer operator to be between the batter's waist and knee (typically 2.5 feet above the ground). That this region varies among batters is not a problem, as all of the analyses we present include batter fixed effects.

Pitches in certain locations are almost always called one way or the other. This is apparent in Figure 4, which shows the location of all called strikes. A strike generally corresponds to the elliptical region centered around the plate and slightly below the batter's waistline. We define three concentric ellipses corresponding to: 1) the *inside* of the strike zone, 2) the *edge* of the strike zone, located just outside the center region, and 3) the complement to both regions, denoted as *outside*. Figure 4 shows the *inside*, an ellipse with major axis equal to 2 feet, and a minor axis equal to 1.6 feet. The *edge* is bordered by the *inside* and the *outside*, a larger ellipse with major axis 2.6 feet and minor axis 2.2 feet. We experimented with several alternative sizes for these ellipses, and none changes the basic results. Panel B of Table 6 summarizes the distribution of pitches by region. Roughly 40 percent are thrown in each of the *inside* and *outside* regions, with the balance in the *edge.*

Pitches thrown to each region generate different outcomes. A called pitch in the *inside* region will be a strike almost 87 percent of the time. Thus, a pitch thrown in this region is associated with little uncertainty. Similarly, a pitch thrown in the *outside* region has very little chance of being called a strike (3.8 percent), resulting in even less uncertainty about the call. A pitch thrown to the *edge* region, however, is called a strike 44.3 percent of the time, generating

nearly the maximum uncertainty possible for a binomial variable. The *edge* region allows the umpire the greatest discretion.

Given this distinction, it is comforting that the *edge* is where the effects of the previous sections occur. Matches in the *inside* are associated with an increase in the called strike percentage of only 0.3 percentage points, from 86.7 percent (no match) to 87 percent (match). The *outside* shows no difference at all. The percent called strikes in the *edge* is 43.6 absent a match, compared to 44.5 percent with a match. If pitchers understand this advantage, then we can predict that a matching pitcher will throw more pitches to the *edge*, where his advantage (courtesy of a biased umpire) is maximized. This aids the pitcher, because pitches to this region are considerably more difficult for the batter to hit.

Panel C of Table 6 presents the results of regressions similar to (1), except: 1) we include *all* pitches thrown by starting pitchers, not just called pitches, as was required for the previous analysis; and 2) the dependent variable indicates whether a pitch is thrown to the *edge*. As before, we include fixed effects for each pitcher, umpire, and batter, as well as all count and inning indicators. The first column shows the result for pitchers in QuesTec parks, where we see that a race/ethnicity match between the pitcher and umpire has virtually no effect on pitch location. In non-QuesTec parks, the situation changes drastically. Matches lead to a 0.95 percentage-point increase in the probability of throwing to the middle region, representing a 5 percent increase relative to the base non-match rate of 19.7 percent. The third column aggregates all observations, where the magnitude of the interaction term is over 1 percent (p=.10). By throwing pitches that can reasonably be called as either balls or strikes, matching pitchers gamble on the fact that this region offers them an advantage.

Panel D of Table 6 shows a related, but distinct, finding. Its interpretation requires some institutional detail. The most common pitch in baseball is the *fastball* (about 58 percent of our sample), which travels in a mostly straight line from the pitcher's hand toward home plate. Skilled pitchers, however, can place spin on pitches, causing them to deviate from a straight trajectory. Pitches with substantial "break" end their flights with dramatic dips that are notoriously difficult to hit solidly. Adding this vertical element also makes these pitches more difficult to judge.[23] As with pitches to the *edge*, judging a curveball requires subjectivity, which is the source of a matching pitcher's advantage. If matching pitchers are aware of a biased umpire, we would expect them to throw more breaking pitches.

The first column in Panel D shows that, in QuesTec parks, a match is associated with a slight preference for breaking balls. In non-QuesTec parks, the magnitude quadruples to 1.28 percent ($p<0.001$). The aggregation of all pitches in Column (3) tells the same story. Matching pitchers in parks without explicit monitoring select pitches allowing umpires the most discretion, enabling them to maximize their advantage stemming from the umpire's bias. While Panel D only makes the distinction between curveballs and other pitches, the result is nearly identical if we distinguish between all breaking pitches (e.g, sliders, cutters) and fastballs.

The Appendix presents a simple game-theoretic model that formalizes the intuition for the results in Table 6. It shows that, when pitchers expect a racial/ethnic match with the umpire to result in more called strikes, their optimal response is to select pitch locations further from the center of the plate (as shown in Table 6). Intuitively, the umpire's bias reduces the penalty for throwing *edge* pitches that are difficult for the batter to hit.

[23]A Google search for "umpire" and "calling a curveball" generates hundreds of links to articles and advice to umpires wishing to master the evaluation of this difficult pitch.

The results and the general theory seem relevant for examining the effect of bias on agents' behavior in a variety of contexts. For example, in the literature on racial profiling (e.g., John Knowles *et al*, 2001; Nicola G. Persico, 2002), while the search data in the empirical literature do not allow examining these indirect effects, the theory demonstrates that they will arise. On the reverse side, the theory of affirmative action (Stephen Coate and Glenn C. Loury, 1993) demonstrates that anti-discriminatory policies will produce indirect effects on agents' behavior.

In the larger labor market the history of occupational segregation is replete with examples of discrimination in occupational choice altering agents' labor-market behavior to their own detriment. The exclusion of Jews from property ownership in the late Middle Ages, the exclusion of African-Americans from most of the railway trades until the 1950s, and perhaps even the "glass ceilings" in corporate hierarchies, all resulted in crowding into occupations (see Barbara R. Bergmann, 1971) that was an indirect effect of bias in other occupations. Our work merely provides a specific example of these effects that allows them to be identified more clearly than in the broader labor-market context.

VI. Measures of Performance and the Measurement of Discrimination

The model in the Appendix and evidence in Section V jointly imply that, conditional on swinging, the batter is less likely to get a hit when the umpire and pitcher match. This implication suggests analyzing a variety of game-level performance measures for each starting pitcher to infer the total of the direct and indirect effects of bias on performance. Table 7 examines each starting pitcher's hits allowed, runs given up, and wins (per game).[24] Because the

[24]We tabulate each starting pitcher's win decisions rather than whether the team actually wins the game. If one considers this second measure instead, the differences are similar, although the overall mean is 0.5 by construction. (The mean for wins is lower in the text table because relief pitchers are frequently awarded decisions.)

sample shrinks by nearly three orders of magnitude compared to the pitch-level results, our ability to detect relatively subtle effects is greatly reduced. Nonetheless, for both groups (non-Hispanic Whites in Panel A, minorities in Panel B), pitchers' outcomes along all three game-level performance measures are superior in matching situations. Non-Hispanic White starting pitchers who match are awarded wins 1.7 percentage points more often in non-QuesTec parks, which reverses to negative 3 percentage points in QuesTec parks. The "QuesTec effect" of 4.6 percentage points is nearly significant (p=.08). For minority starting pitchers, the similar gap is even larger, at 12.9 percentage points (p=.06), although there are only 74 matches in QuesTec parks.

Several other aggregate performance measures show the same patterns. Both groups give up fewer hits in matching situations in non-QuesTec parks, Whites by about 1 percentage point, minorities by about 2 percentage points. As before, each pattern reverses in QuesTec parks. A similar pattern is seen along additional performance metrics. Figure 5 shows several of them, again for non-Questec parks and for White and minority pitchers separately. Presented as percentage changes from their baseline levels (Table 7 presented differences in levels), the vast majority improve in match situations. From the starting pitcher's perspective, a racial/ethnic match with the umpire helps his earned runs (fewer), hits (fewer), walks (fewer), and home runs (fewer). Only strikeouts go in the opposite direction. One might expect little effect for strikeouts, which, at least in the fraction that are called third strikes, require bias on a terminal count, which we have already shown does not occur.

To the extent that pay is based on measured productivity, our findings of small direct and larger indirect effects of racially/ethnically disparate treatment carry important implications for measuring the extent of discrimination in baseball and in labor markets generally. In particular,

they imply that estimates of the extent of discrimination will be understated, even controlling for standard measures of performance.

Consider a simple earnings equation:

$$W_i = \alpha M_i + \beta P^*_i + \upsilon_i , \qquad (2)$$

where W is the logarithm of earnings, M an indicator of minority status, P* worker i's true productivity, and υ a random error in the determination of earnings. The parameter α is the true effect of minority status on earnings when productivity measurements are free of bias. Assume that the majority workers' productivity is measured without bias, but that minority workers are subject to a negative bias in their assessment by evaluators, which leads to a shortfall of their measured productivity P below their true productivity:

$$P_i = P^*_i - \varphi, \text{ if } M=1; \qquad (3)$$

$$P_i = P^*_i , \quad \text{if } M=0,$$

φ>0. Then we can rewrite (2) to obtain an estimating equation in observables:

$$W_i = [\alpha+\beta\varphi]M_i + \beta P_i + \upsilon_i , \text{ or} \qquad (2')$$

$$W_i = \alpha' M_i + \beta P_i + \upsilon_i .$$

The standard estimate of earnings discrimination adjusted for productivity differences, α', has a positive bias in the amount βφ.

To obtain some feel for the size of this bias in the particular case that we have examined, we can simulate the wage effects using the estimates of φ underlying Figure 5 and estimates of β from three studies of MLB that examined pitchers and used at least some of these outcomes as determinants of salaries. We are essentially estimating the reduction in minority pitchers' salaries as a result of the average amount of bias arising during the 2004-2008 seasons due to umpire-pitcher racial/ethnic matches. Lawrence M. Kahn (1993, Table A2) estimates equations

like (2') using a set of outcome measures that can be conformed to ours by including the percentage of games won and ERA. Making reasonable assumptions about the means of these outcomes for starting pitchers in 2004-2008, applying the effects in Figure 5, and using his parameter estimates yields an estimated bias of $\beta\varphi = 0.034$. Mark P. Gius and Timothy P. Hylan (1996, Table 6.2) use strikeouts/inning, walks/inning and winning percentage, all of which are also conformable with our outcome measures. The same method based on their parameter estimates produces an estimate of $\beta\varphi = 0.012$. Finally, using the estimates for starting pitchers by Anthony C. Krautmann *et al* (2003), the estimate of $\beta\varphi = 0.074$.[25]

While we have demonstrated the extent of bias to estimated discrimination in earnings that arises because of biased evaluations of MLB pitchers, this effect is probably smaller than would be observed for workers generally. The scope for the expression of racial/ethnic preferences of umpires for/against pitchers is almost surely far less than in most workplaces. Evaluations of pitchers are made discretely and very frequently—when a pitch is thrown. These are not one-shot comments made at most monthly at the evaluator's leisure. Also, as our demonstrations of reduced bias when there is greater scrutiny suggest, there are quite stringent external limits on the expression of bias against unmatched pitchers. The relative lack of such limits in the general workplace suggests that the example here may provide a lower bound on the extent of bias to estimates of disparate outcomes generally.

The general point, that bias will affect measures of productivity, is not new (see, e.g., Glen G. Cain, 1986). It is, however, generally ignored in the scholarly literature measuring the

[25]We use the means in this sample as the baselines. For the fraction of games won, 0.37; for ERA, 4.44; for strikeouts/inning by starting pitchers, 0.75; and for walks/inning, 0.43. We can take the estimates of the bias as examples here to infer the dollar impacts of this subtle form of discrimination. In 2006, the mid-point of our sample, the average salary of starting pitchers in MLB was $4.8 million. A bias to the estimated effect of minority status on compensation of starting pitchers of between 1 and 8 percent suggests that those pitchers are underpaid relative to White pitchers by between $50,000 and $400,000 per year.

wage effects of discrimination. In the huge industry of employment litigation, standard practice is to adjust wages using measures of supervisors' evaluations of workers. As we have shown, even in a very controlled and highly scrutinized environment, these can be biased against minorities. Our results suggest that this bias must be accounted for whenever one wishes to measure racial/ethnic disparities in rewards in the workplace.

VII. Conclusions

The analyses of individual pitches and game outcomes suggest that baseball umpires express racial/ethnic preferences in their decisions about players' performances. Pitches are slightly more likely to be called strikes when the umpire shares the race/ethnicity of the starting pitcher, an effect that only is observable when umpires' behavior is not well monitored. The evidence also suggests that this bias has substantial effects on pitchers' measured performance and games' outcomes. The link between the small and large effects arises, at least in part, because pitchers alter their behavior in potentially discriminatory situations in ways that ordinarily would disadvantage themselves (such as throwing pitches directly over the plate). As in many other fields, racial/ethnic preferences work in all directions—most people give preference to members of their own group. In MLB, as in so many other fields of endeavor, power belongs disproportionately to members of the majority—White—group.

The type of discrimination that we have demonstrated is disturbing because of its implications for the sports labor market. In particular, minority pitchers are at a significant disadvantage relative to their White peers, *even in the absence of explicit wage discrimination by teams*. Although some evidence suggests such explicit discrimination exists, i.e., there is a wage gap among baseball players of different races, the fact that almost 90 percent of the umpires are

White implies that the *measured* productivity of minority pitchers may be downward biased. Implicitly, estimates of wage discrimination in baseball that hold measured productivity (at least of pitchers) constant will understate its true size.

More generally, our results suggest caution in interpreting any estimates of wage discrimination stemming from equations relating earnings to race/ethnicity, even with a large set of variables designed to control for differences in productivity. To the extent that supervisors' evaluations are among the control variables included in estimates of wage discrimination, or even if they only indirectly alter workers' objective performances, their inclusion or their mere existence contaminates attempts to infer discrimination from adjusted racial/ethnic differences in wages. If racial/ethnic preferences in evaluator-worker matches are important, standard econometric estimates will generally understate the magnitude of racial/ethnic discrimination in labor markets.

While the specific evidence of racial/ethnic match preferences is disturbing, our novel analysis of the expression of discrimination should be encouraging: When their decisions matter more, and when evaluators are themselves more likely to be evaluated by others, our results suggest that these preferences no longer manifest themselves. Indeed, these findings imply that the particular impacts of racial/ethnic match preferences in baseball may now have been vitiated, since beginning in 2009 all ballparks are equipped with QuesTec or similar technologies.[26] Clearly, however, raising the price of discrimination in the labor market generally is more difficult; but our results may suggest analogous measures that might have the desired effects.

[26] See "Ball-Strike Monitor May Reopen Wounds" (Alan Schwarz, *New York Times*, March 1, 2009, electronic version available at http://www.nytimes.com/2009/04/01/sports/baseball/01umpires.html.

REFERENCES

Blair, Irene V. "The Malleability of Automatic Stereotypes and Prejudice." *Personality and Social Psychology Review*, 2002, *6*(3), pp. 242-61.

Bergmann, Barbara R. "The Effect on White Incomes of Discrimination in Employment." *Journal of Political Economy*, 1971, *79(*2), pp. 294-313.

Bradbury, J.C. *The Baseball Economist: The Real Game Exposed*. New York: Dutton, 2007.

Cain, Glen G. "The Economic Analysis of Labor Market Discrimination: A Survey." In Ashenfelter, Orley C. and Layard, P. Richard G. *Handbook of Labor Economics, Vol. 1*. Amsterdam: North-Holland, 1986.

Chen, Jowei. "What Does Baseball Teach Us About Reducing Racial Discrimination? Evidence From Two Natural Experiments." Stanford University working paper, 2007.

Coate, Stephen and Loury, Glenn C. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 1993, *83*(5), pp. 1220-40.

Dohmen, Thomas J. "The Influence of Social Forces; Evidence from the Behavior of Football Referees." *Economic Inquiry*, 2008, *46*(3), pp. 411-24.

Donald, Stephen G. and Hamermesh, Daniel S."What Is Discrimination? Gender in the American Economic Association." *American Economic Review*, 2006, *96*(4), pp. 1283-92.

Findlay, David W. and Reid, Clifford E. "Voting Behavior, Discrimination and the National Baseball Hall of Fame." *Economic Inquiry*, 1997, *35*(3), pp. 562-78.

Fort, Rodney D. and Gill, Andrew M. "Race and Ethnicity Assessment in Baseball Card Markets." *Journal of Sports Economics*, 2000, *1*(1), pp. 21-38.

Garicano, Luis; Palacios-Huerta, Ignacio and Prendergast, Canice. "Favoritism under Social Pressure." *Review of Economics and Statistics*, 2005, *87*(2), pp. 208-16.

Gius, Mark P. and Hylan, Timothy R. "An Interperiod Analysis of the Salary Impact of Structural Changes in Major League Baseball: Evidence from Panel Data." In *Baseball Economics, Current Research*. Westport CT: Greenwood, Praeger, 1996.

Goodwin, Stephanie A.; Gubin, Alexandra; Fiske, Susan T. and Yzerbyt, Vincent Y. "Power Can Bias Impression Processes: Stereotyping Subordinates by Default and By Design." *Group Processes and Intergroup Relations*, 2000, *3*(3), pp. 227-56.

Gwartney, James D. and Haworth, Charles T. "Employer Costs and Discrimination: The Case of Baseball." *Journal of Political Economy*, 1974, *82*(4), pp. 873-81.

Kahn, Lawrence M. "Racial Discrimination in Professional Sports: A Survey of the Literature." *Industrial and Labor Relations Review*, 1991, *44*(3), pp. 395-418.

----------, "Managerial Quality, Team Success and Individual Player Performance in Major League Baseball." *Industrial and Labor Relations Review*, 1993, *46*(3), pp. 531-47.

Knowles, John; Persico, Nicola G. and Todd, Petra E.. "Racial Bias in Motor-Vehicle Searches: Theory and Evidence." *Journal of Political Economy*, 2001, *109*(1), pp. 203-29.

Krautmann, Anthony C.; Gustafson, Elizabeth F. and Hadley, Lawrence. "A Note on the Structural Stability of Salary Equations: Major League Baseball Pitchers." *Journal of Sports Economics*, 2003, *4*(1), pp. 56-63.

Nardinelli, Clark and Simon, Curtis J. "Customer Racial Discrimination in the Market for Memorabilia: The Case of Baseball." *Quarterly Journal of Economics*, 1990, *105*(3), pp. 575-95.

Pascal, Anthony H. and Rapping, Leonard A. "The Economics of Racial Discrimination in Organized Baseball." In Pascal, Anthony H., ed., *Racial Discrimination in Economic Life*. Lexington, MA: Heath, 1972.

Persico, Nicola G. "Racial Profiling, Fairness, and Effectiveness of Policing." *American Economic Review*, 2002, *92*(5), pp. 1472-97.

Price, Joseph and Wolfers, Justin J.. "Racial Discrimination Among NBA Referees." National Bureau of Economic Research, Working Paper No. 13206, June 2007.

Scully, Gerald W.. "Pay and Performance in Major League Baseball." *American Economic Review,* 1974*, 64*(6), pp. 915-30.

Stoll, Michael A.; Raphael, Steven and Holzer, Harry J. "Black Job Applicants and the Hiring Officer's Race." *Industrial and Labor Relations Review*, 2004, *57*(2), pp. 267-87.

Zitzewitz, Eric W. "Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making." *Journal of Economics and Management Strategy*, 2006, *15*(1), pp. 67-99.

Table 1–Summary Statistics of Pitches

Pitch Outcomes, 2004-2008 (percentage distributions)								
	Total Pitches	**Called Strike**	**Called Ball**	**Swinging Strike**	**Foul**	**In Play**	**Intentional Ball**	**Hit by Pitch**
All	3,524,624	17.09	36.56	8.98	17.08	19.41	0.63	0.25
Pitcher								
White (N=861)	2,544,515	17.19	36.48	8.77	17.10	19.58	0.64	0.24
Hispanic (N=278)	793,797	16.86	36.77	9.57	17.03	18.86	0.64	0.27
Black (N=37)	89,355	16.24	36.68	9.71	17.54	19.07	0.52	0.24
Asian (N=39)	96,957	17.12	36.81	8.87	16.59	19.70	0.64	0.27
Batter								
White (N=1,147)	1,813,768	17.37	36.90	9.11	16.92	18.84	0.58	0.28
Hispanic (N=493)	1,061,115	16.81	35.91	8.72	17.35	20.31	0.68	0.22
Black (N=187)	571,563	16.65	36.67	9.17	17.08	19.50	0.70	0.23
Asian (N=46)	78,178	17.63	36.80	7.44	17.35	19.88	0.71	0.19
Umpire								
White (N=91)	3,215,949	17.09	36.55	8.97	17.09	19.41	0.64	0.25
Hispanic (N=5)	111,524	17.06	36.80	8.87	16.97	19.33	0.70	0.27
Black (N=6)	197,151	17.13	36.63	9.00	16.99	19.44	0.59	0.22

Table 2–Summary of Umpires' Called Pitches by Umpire-Pitcher Racial/Ethnic Match, MLB 2004-2008

	Pitcher Race/Ethnicity				Total percent called strikes
	White	Hispanic	Black	Asian	
Umpire Race/Ethnicity					
White					
Pitches	2,319,522	726,137	81,251	89,039	
Called pitches	1,244,523	389,411	42,986	47,973	
Called strikes	398,673	122,441	13,194	15,269	
Percent called strikes	32.03	31.44	30.69	31.83	31.86
Hispanic					
Pitches	80,956	24,844	2,559	3,165	
Called pitches	43,632	13,299	1,374	1,760	
Called strikes	13,857	4,194	429	549	
Percent called strikes	31.76	31.54	31.22	31.19	31.68
Black					
Pitches	144,037	42,816	5,545	4,753	
Called pitches	77,472	23,035	2,922	2,561	
Called strikes	24,900	7,195	886	784	
Percent called strikes	32.14	31.24	30.32	30.61	31.86
Total percent called strikes	32.03	31.43	30.69	31.75	31.86

Table 3–Effects on Called Strikes of the Relationship between Pitcher and Umpire Race/Ethnicity, MLB 2004-2008

A. Main Parameter Estimates

Pitchers	***White***	***Black***	***Hispanic***	***All***	***All***	***All***	***All***	***All***	***All***
Umpires	***All***	***All***	***All***	***White***	***Black***	***Hispanic***	***All***	***All***	***All***
	(1)	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***	***(7)***	***(8)***	***(9)***
Black Umpire	-0.0005	0.0004	-0.0010						
	(0.0019)	(0.0105)	(0.0031)						
Hispanic Umpire	-0.0045	0.0097	0.0079						
	(0.0024)	(0.0127)	(0.0049)						
Black Pitcher				-0.0148	-0.0157	-0.0027			
				(0.0023)	(0.0103)	(0.0125)			
Hispanic Pitcher				-0.0072	-0.0089	0.0020			
				(0.0009)	(0.0034)	(0.0054)			
UPM							0.0024	0.0021	0.0016
							(0.0013)	(0.0017)	(0.0017)
Observations	1,365,660	47,285	425,731	1,676,942	103,429	58,305		1,838,676	
R^2	0.031	0.031	0.030	0.028	0.025	0.030	0.031	0.091	0.091
Fixed Effects	P	P	P	U	U	U	P	PU	PUB

B. Coefficients on Pitch Count Indicators in the Specification in Column (9)

0&1	***0&2***	***1&0***	***1&1***	***1&2***	***2&0***	***2&1***	***2&2***	***3&0***	***3&1***	***3&2***
-0.226	-0.355	-0.023	-0.190	-0.326	0.042	-0.151	-0.289	0.206	-0.060	-0.257
(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)

C. Coefficients on Inning Indicators and Pitcher's Score Advantage in the Specification in Col. (9)

2^{nd}	***3^{rd}***	***4^{th}***	***5^{th}***	***6^{th}***	***7^{th}***	***8^{th}***	***9^{th}***	***Top of Inning***	***Pitcher's Score Advantage***
-0.010	-0.024	-0.032	-0.032	-0.034	-0.025	-0.024	-0.018	0.006	0.002
(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.0004)

Notes: All estimates are based on linear-probability models with heteroskedasticity-robust standard errors in parentheses, here and in Tables 4-6. UPM indicates whether the umpire and pitcher match on race/ethnicity. The control variables whose coefficients are reported Panels B and C are included in all the estimates. Pitcher Score Advantage is the number of runs, potentially negative, that the pitcher's team is ahead at the time of the pitch. Top of Inning is an indicator equaling 1 if the home team is pitching. P, U and B represent pitcher, umpire and batter fixed effects, respectively.

Table 4–Effects on Called Strikes of Explicit Monitoring of Umpires and Racial/Ethnic Discrimination, MLB 2004-2008

	QuesTec	*Non-QuesTec*	*All*
	(1)	*(2)*	*(3)*
Umpire-Pitcher Match (UPM)	-0.0048	0.0059	0.0059
	(0.0027)	(0.0022)	(0.0022)
QuesTec*UPM			-0.0107
			(0.0035)
Observations	679,979	1,158,697	1,838,676
R^2	0.089	0.088	0.088

Notes: UPM indicates whether the umpire and pitcher match on race/ethnicity. All columns here and in Tables 5 and 6 include the control variables shown in Table 3. All columns also include fixed effects: 1) pitcher-QuesTec fixed effects, i.e., two fixed effects for each pitcher who pitched in both a ballpark where QuesTec was and was not installed; 2) umpire-QuesTec fixed effects, and 3) batter-QuesTec fixed effects.

Table 5–Effects on Called Strikes of Implicit Monitoring of Umpires and Racial Discrimination, MLB 2004-2008

A. Distinguishing by Game Attendance

	High Attendance *(1)*	*Low Attendance* *(2)*	*All Games* *(3)*
UPM	-0.0034	0.0064	0.0036
	(0.0025)	(0.0024)	(0.0019)
Well Attended (>69% capacity)			0.0059
			(0.0012)
Well Attended*UPM			-0.0037
			(0.0015)
Observations	902,261	936,415	1,838,676
R^2	0.089	0.088	0.088

B. Distinguishing by Terminal Count and Inning

	Terminal *(4)*	*Non-Terminal* *(5)*	*All Pitches* *(6)*	*Early Inning* *(7)*	*Late Inning* *(8)*
UPM	-0.0026	0.0031	0.0031	0.0044	0.0023
	(0.0027)	(0.0021)	(0.0018)	(0.0031)	(0.0022)
Terminal Count			-0.0058	-0.0086	-0.0038
***UPM**			(0.0014)	(0.0024)	(0.0017)
Observations	427,136	1,411,540	1,838,676	641,053	1,197,623
R^2	0.175	0.042	0.088	0.095	0.085

C. Combining Explicit and Implicit Monitoring Proxies

		UPM Interacted with				
	UPM	**Questec**	**Well Attended**	**Terminal Count**	**Observations**	**R^2**
All Pitches	0.0089	-0.0102	-0.0035	-0.0058	1,838,676	0.088
(9)	(0.0024)	(0.0036)	(0.0015)	(0.0014)		

Note: Low (high) attendance games are defined as games with percentage attendance below (above) the median. A terminal count is defined as a count with three balls and/or two strikes.

Table 6–Pitch Location, Type and the Effects of Pitcher-Umpire Racial/Ethnic Matches, MLB 2007-2008

A. Pitch Locations (Distance from Home-plate Center), 2007-2008, N=538,194

	Mean	Quantiles:	5th	25th	50th	75th	95th
Horizontal Pitch Distance (ft.)	-0.04		-1.53	-0.68	-0.04	0.60	1.44
Pitch Height (ft.)	-0.11		-1.60	-0.69	-0.10	0.48	1.36

B. Percentage Distributions of Pitches by Type, 2007-2008, N=533,150

By Location			*By Type*				
Inside	***Edge***	***Outside***	***Change-up***	***Curveball***	***Fastball***	***Slider***	***Other***
39.55	19.98	40.47	13.43	10.88	57.48	13.52	4.69

C. Effects on Probability of Pitch in the Edge of the Strike Zone

	Questec (1)	*Non-Questec (2)*	*All Games (3)*
UPM	-0.0005	0.0095	0.0095
	(0.0047)	(0.0042)	(0.0042)
Questec*UPM			-0.0102
			(0.0063)
Observations	199,085	339,109	538,194
R^2	0.001	0.001	0.001

D. Effects on Probability of a Curve Ball

	Questec (4)	*Non-Questec (5)*	*All Games (6)*
UPM	0.0033	0.0128	0.0125
	(0.0032)	(0.0029)	(0.0028)
Questec*UPM			-0.0087
			(0.0043)
Observations	195,777	337,373	533,150
R^2	0.020	0.021	0.021

Notes: The sample consists of all pitches (called and non-called, excluding intentional balls) thrown by starting pitchers In Panel A the pitch location is the Cartesian coordinate, where the origin is the intersection of the vertical line from the center of the home plate and the horizontal line equidistant to the top and the bottom of the strike zone. The information is from PITCHf/x.

Table 7–Estimated Effects on Performance of Umpire and Starting Pitcher Racial/Ethnic Match, N=12,127 Games, MLB 2004-2008

A. White Pitchers

	Umpire-Pitcher Racial Match	*N*	*Win*	*Hits Allowed*	*Runs Allowed*
Questec	Match	5,953	0.347	6.190	3.215
	Non-Match	605	0.377	6.109	3.179
	Diff		-0.030	0.081	0.036
			(0.021)	(0.102)	(0.092)
Non-Questec	Match	10,491	0.351	6.174	3.154
	Non-Match	1,003	0.334	6.240	3.234
	Diff		0.017	-0.066	-0.080
			(0.016)	(0.073)	(0.069)
	Diff-in-Diff		-0.046	0.147	0.116
			(0.026)	(0.126)	(0.115)

B. Minority Pitchers

	Umpire-Pitcher Racial Match	*N*	*Win*	*Hits Allowed*	*Runs Allowed*
Questec	Match	74	0.257	6.284	3.581
	Non-Match	2,313	0.356	6.006	3.179
	Diff		-0.099	0.278	0.402
			(0.052)	(0.297)	(0.276)
Non-Questec	Match	119	0.370	5.891	3.185
	Non-Match	3,696	0.340	6.080	3.223
	Diff		0.030	-0.189	-0.038
			(0.045)	(0.226)	(0.214)
	Diff-in-Diff		-0.129	0.466	0.440
			(0.069)	(0.373)	(0.349)

Notes: Standard errors in parentheses.

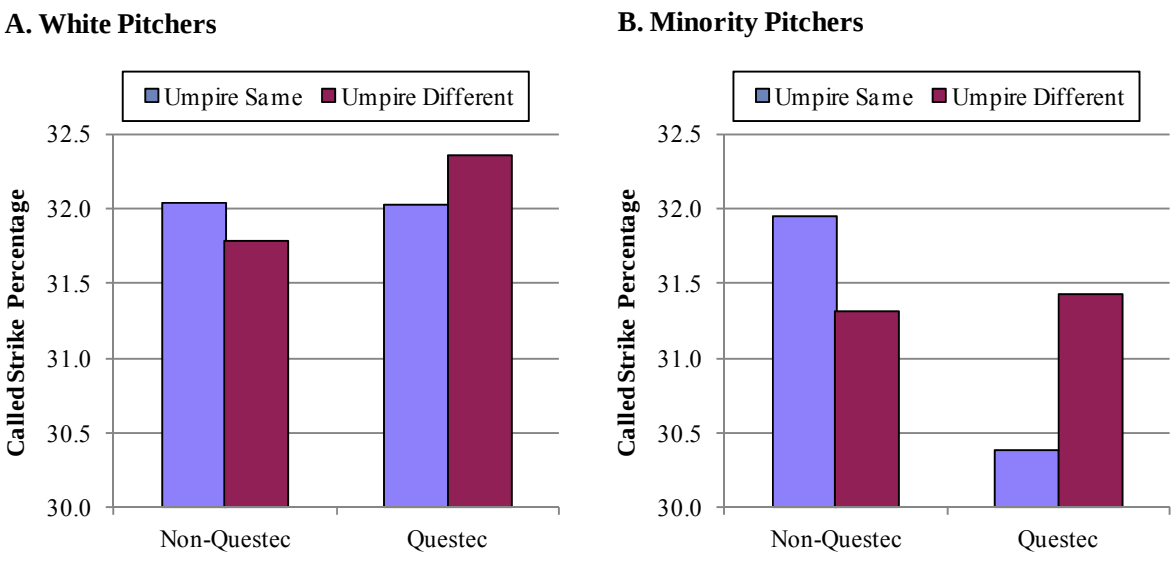

Figure 1–Race and Called Strike Percentages in QuesTec and Non-QuesTec Ballparks

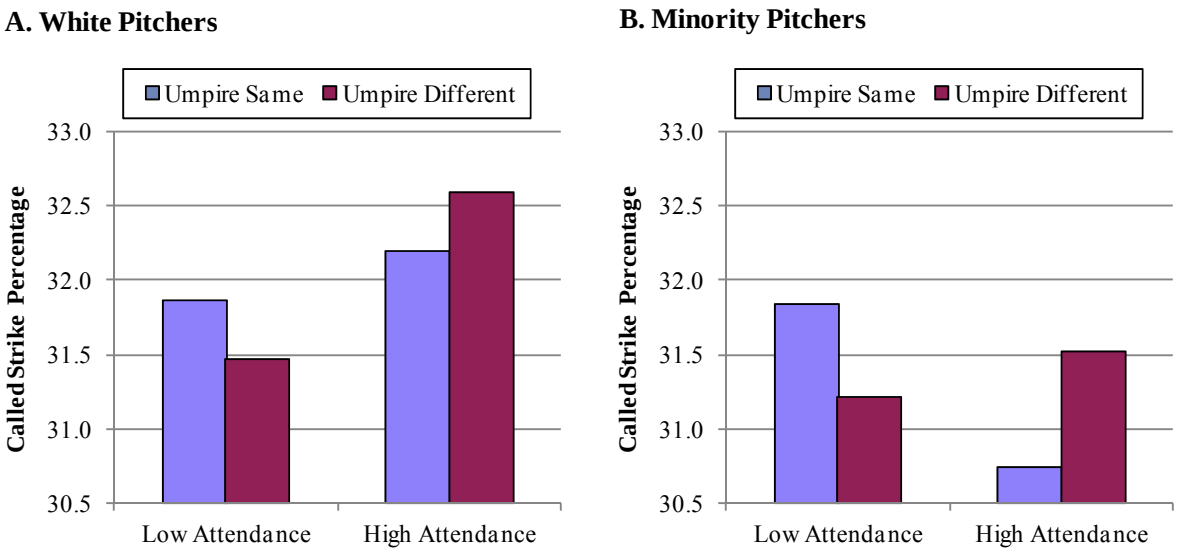

Figure 2–Race and Called Strike Percentages by Game Attendance

Note: Low (high) attendance games are defined as games with percentage attendance below (above) the median.

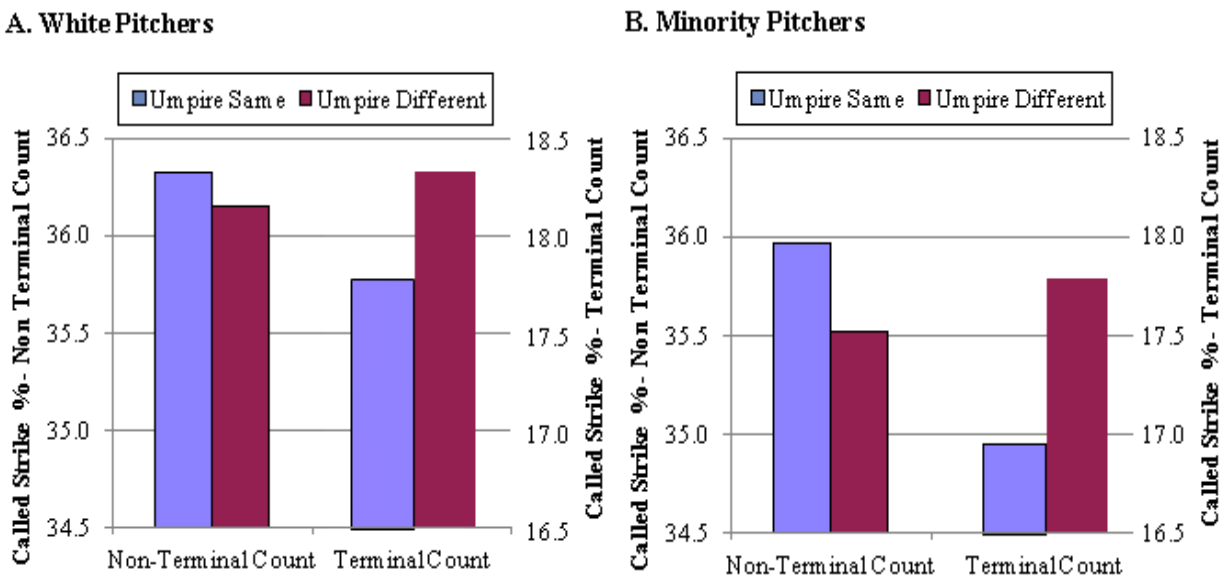

Figure 3–Race and Called Strike Percentages in Terminal and Non-Terminal Counts

Note: A terminal count is defined as a count with three balls and/or two strikes.

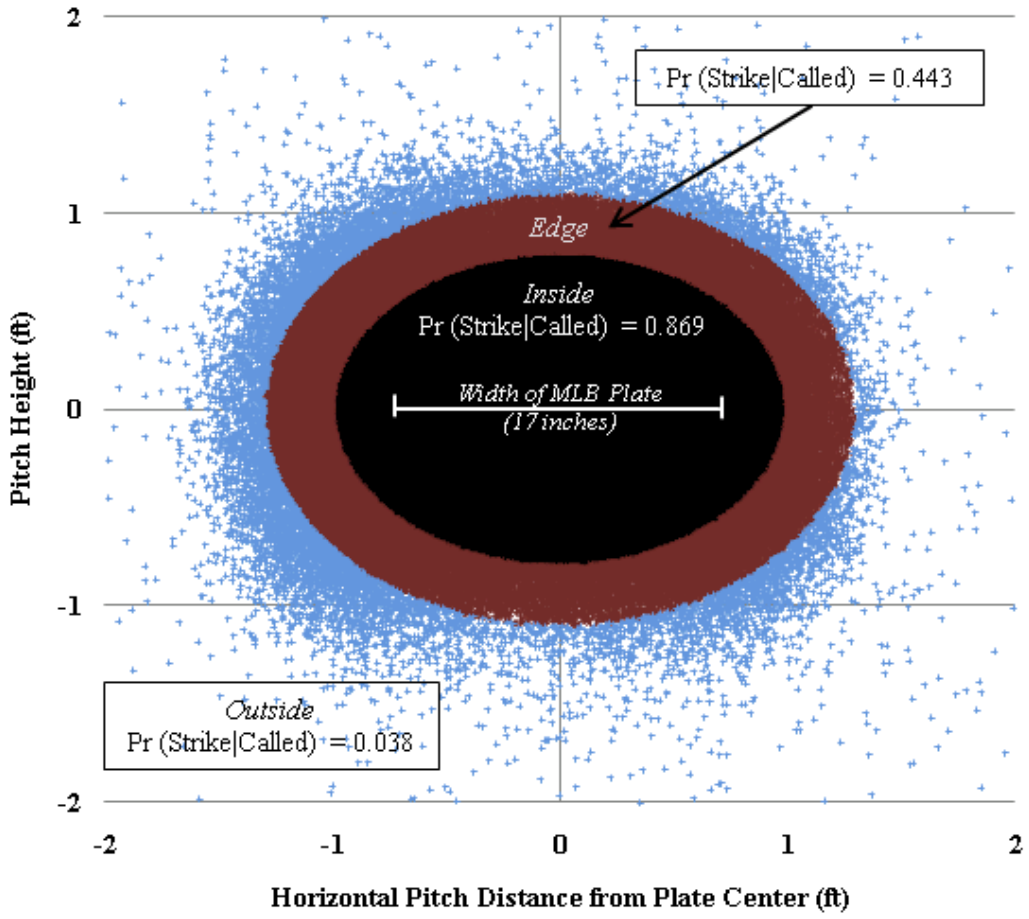

Figure 4–Called Strikes by Distance from Home-Plate Center, 2007-2008 (N=144,990)

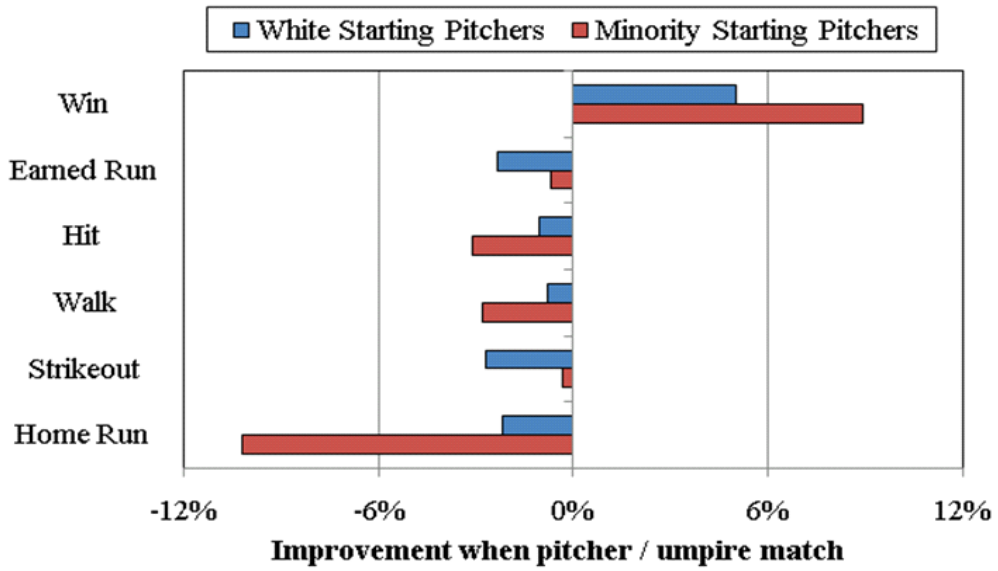

Figure 5–Effects of Umpire-Pitcher Racial/Ethnic Match on Pitcher Performance, non-Questec Ballparks, MLB 2004-2008 (N=15,308)

Appendix– A Model of Bias-Induced Changes in Pitcher Strategies

Consider the following simple representation of the interaction between the pitcher and hitter. Denote the horizontal distance from the center of the plate $\tilde{y}$. Assume for simplicity that the pitcher can control the width of pitches (i.e., the horizontal dimension), but not their height. Further suppose that the batter is left-handed, and that the pitcher never aims left of center, i.e., $\mu = 0$. This simplifying assumption is little more than a normalization, although a realistic one, as pitchers are usually cautious to avoid hitting the batter.

The game unravels as follows.

1. The pitcher moves first. He can select his aim, $\mu \geq 0$, but not the final pitch location, $\tilde{y}$, which is random. On average, the pitcher's aim is correct, i.e., $E(\tilde{y}) = \mu$.
2. The batter moves next. A batter must decide whether to swing or not soon after a pitch is thrown, but before it reaches its final location $\tilde{y}$. To capture this timing, the batter's swing decision is made immediately after observing μ.[27]
3. If the batter does not swing, two outcomes are possible. For a given value of μ, with probability $s(\mu)$, the pitch is called a strike, and confers the batter a payoff S. With probability $1 - s(\mu)$, the pitch is called a ball, with payoff $B > S$**.** We assume $s' < 0$, $s'' < 0$, i.e., that pitches aimed closer to the plate are more likely to be called strikes, and at an increasing rate.
4. If the batter swings, two additional outcomes are possible. With probability $h(\mu)$, the batter gets a hit, and enjoys a payoff H. With probability $1 - h(\mu)$, the batter does not get a hit, with payoff $N < H$.[28] Similar to the assumptions for s, we assume $h' < 0$, $h'' < 0$**.**

The Batter's Problem:

To determine whether he swings at a pitch with expected location μ, the batter compares his expected payoff from swinging,

$$\pi(swing|\mu) = h(\mu)H + [1 - h(\mu)]N \qquad (1)$$

[27] This strict timing assumption is not crucial. Instead, it is a simplified way of modeling that the batter makes his swing decision under imperfect information. For example, the batter could instead observe a noisy signal of $\tilde{y}$ without changing the results.

[28] Here, N captures the average payoff of swinging and missing, (S), and hitting into an out.

with that from not swinging,

$$\pi(\text{no swing}|\mu) = s(\mu)S + [1 - s(\mu)]B. \quad (2)$$

Lemma 1. *Assume* $\pi(\text{swing} \mid \mu = 0) > \pi(\text{no swing}) \mid \mu = 0)$, *so that a batter always prefers to swing at a pitch aimed down the center of the plate. Then there exists a unique cutoff M whereby if: i)* $\mu < \hat{\mu}$, *the batter strictly prefers to swing, ii)* $\mu > \hat{\mu}$, *the batter strictly prefers to not swing, and iii)* $\mu = \hat{\mu}$, *the batter is indifferent between swinging and not.*

Proof. *$\partial(\pi(\text{swing}|\mu)) / \partial\mu < 0$, which follows from the assumptions that: i) called strikes are assumed to be more likely when thrown closer to the plate, $s' < 0$, ii) the batter's expected payoff from called balls is higher than that from called strikes, $B > S$. By similar logic, $\partial(\pi(\text{no swing}|\mu)) / \partial\mu > 0$. The convexity assumptions $s'' < 0$, $h'' < 0$ then guarantee a single crossing for $(\pi(\text{swing})|\mu)$ and $(\pi(\text{no swing})|\mu)$, which we denote* $\hat{\mu}$.

The intuition for Lemma 1 is straightforward. Batters will not attempt to hit pitches that have very little chance of being called a strike should they not swing, i.e., for sufficiently low values of μ. Moreover, the cutoff for swinging $\hat{\mu}$ is a function of the payoffs *S, B, H*, and *N* that correspond to the possible outcomes of the plate appearance. Generally, these payoffs will depend on game conditions, such as the score, the count, runners on base, or other factors that determine the payoffs to each outcome. For example, with runners on second and third base but no outs, the benefit of a hit (*H*) is substantial, where the cost of hitting into an out (*N*) is relatively small. In this situation, the batter will be less selective at the plate, which increases the swinging cutoff $\hat{\mu}$. We do not model differences in these payoffs across plate appearances, although the present set-up easily allows for this extension.

Our main interest is in how changes in the conditional strike function, $s(\mu)$, influence the batter's optimal behavior. Specifically, assume that the race/ethnicity match of the umpire and pitcher influences the probability that a pitch aimed at location μ will be called a strike. If the pitcher and umpire match (*M*), denote the conditional called strike probability $s_M(\mu)$. If they are different (*D*), the conditional strike probability becomes $s_D(\mu)$. To capture the idea that similar race or ethnicity helps the pitcher, we assume:

$$s_M(\mu) > s_D(\mu), \ \forall\mu \quad (3)$$

In other words, the same pitch has a different probability of being called a strike, conditional on whether the umpire and pitcher have the same or different races or ethnicities.

Lemma 2. *When the pitcher and umpire share the same race/ethnicity, the batter swings at pitches further from the center of the plate. That is, the cutoff location under a match is strictly greater than the cutoff location otherwise, i.e.,* $\hat{\mu}_M > \hat{\mu}_D$.

Proof. *Denote* $\hat{\mu}_M$ *as the cutoff swinging location when* $s(\mu) = s_M(\mu)$ *and* $\hat{\mu}_D$ *as that when* $s(\mu) = s_D(\mu)$. *Suppose* $s(\mu) = s_M(\mu)$ *and* $\mu = \hat{\mu}_M$. *From equation (2), when* $s(\mu)$ *changes to* $s_D(\mu)$ *the expected payoff of not swinging declines by* $[s_M(\hat{\mu}_M) - s_D(\hat{\mu}_M)](S - B) > 0$, *while the payoff from swinging is unchanged. We can now use the proof for Lemma 1. Because* $\partial(\pi(swing \mid \mu)) / \partial\mu < 0$ *and* $\partial(\pi(no\ swing \mid \mu)) / \partial\mu > 0$, *the new cutoff* $\hat{\mu}_D$ *is strictly less than* $\hat{\mu}_M$.

Lemma 2 indicates that when the batter anticipates judgments that favor the pitcher, his optimal strategy changes. Expecting the umpire's bias to reduce his payoff from not swinging, the batter takes matters into his own hands by swinging at pitches that he would otherwise let pass. Empirically, this implies a distinct advantage to the pitcher, not only for pitches that are called, but also for pitches that are hit. We complete this exercise by extending consideration to the pitcher's optimal strategy.

The Pitcher's Problem:

The pitcher's choice variable is μ, the expected location of the pitch. His expected payoff is the inverse of the batter's. If the batter swings, then the pitcher's expected payoff is $-h(\mu)H-[1-h(\mu)]N$. If the batter does not swing, then his expected payoff is $-s(\mu)S-[1-s(\mu)]B$.

Lemma 3. *The pitcher's optimal pitch location is* $\hat{\mu}$, *so that the batter is indifferent between swinging and not.*

Proof. *The batter will swing at any pitch aimed at* $\mu < \hat{\mu}$, *but because* $\partial(\pi(swing \mid \mu)) / \partial\mu < 0$, *the pitcher is always strictly better off increasing* μ *given that the batter will swing. The batter will not swing at any pitch aimed at* $\mu > \hat{\mu}$, *but because* $\partial(\pi(no\ swing \mid \mu)) / \partial\mu > 0$, *the pitcher will always decrease* μ *given that the batter will not swing. It follows then that the optimal pitch location must be* $\hat{\mu}$.

The main prediction is that the umpire's bias influences not only called strikes and balls, but also pitches where the umpire's judgment plays no direct role. Lemma 3 shows that the umpire's judgment influences the choice of pitch location, which in turn influences the batter's incentive to swing at the ball. It follows that *conditional on swinging*, the batter is less likely to hit the ball when the umpire and pitcher share race or ethnicity. As indicated by the model, this is because pitches are, on average, more difficult to hit in these situations.

Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians

By Claudia Goldin and Cecilia Rouse*

A change in the audition procedures of symphony orchestras—adoption of "blind" auditions with a "screen" to conceal the candidate's identity from the jury—provides a test for sex-biased hiring. Using data from actual auditions, in an individual fixed-effects framework, we find that the screen increases the probability a woman will be advanced and hired. Although some of our estimates have large standard errors and there is one persistent effect in the opposite direction, the weight of the evidence suggests that the blind audition procedure fostered impartiality in hiring and increased the proportion women in symphony orchestras. (*JEL* J7, J16)

Sex-biased hiring has been alleged for many occupations but is extremely difficult to prove. The empirical literature on discrimination, deriving from the seminal contributions of Gary Becker (1971) and Kenneth Arrow (1973), has focused mainly on disparities in earnings between groups (e.g., males and females), given differences in observable productivity-altering characteristics. With the exception of various audit studies (e.g., Genevieve Kenney and Douglas A. Wissoker, 1994; David Neumark et al., 1996) and others, few researchers have been able to address directly the issue of bias in hiring practices.[1] A change in the way symphony orchestras recruit musicians provides an unusual way to test for sex-biased hiring.

Until recently, the great symphony orchestras in the United States consisted of members who were largely handpicked by the music director. Although virtually all had auditioned for the position, most of the contenders would have been the (male) students of a select

* Goldin: Department of Economics, Harvard University, Cambridge, MA 02183; Rouse: Woodrow Wilson School, Princeton University, Princeton, NJ 08544. Rouse acknowledges The National Academy of Education, the NAE Spencer Postdoctoral Fellowship Program, and the Mellon Foundation for financial support. We are indebted to the staff members of the orchestras that gave us access to their audition records and who provided other assistance, and to the musicians who responded to our questionnaire. We are particularly grateful to Joanne Berry, Brigit Carr, Ruth DeSarno, Stefanie Dyson, Josh Feldman, Barbara Haws, Oren Howard, Cindy Hubbard, Carol Jacobs, Lynn Larsen, Bennett McClellan, Stephen Molina, Bill Moyer, Jeffrey Neville, Stephen Novak, Deborah Oberschalp, Stacey Pelinka, Carl Schiebler, Alison Scott-Williams, Robert Sirineck, Harold Steiman, and Brenda Nelson Strauss. We also thank Gretchen Jackson of the University of Michigan School of Music. Rashid Alvi, Brigit Chen, Eric Hilfers, Serena Mayeri, LaShawn Richburg, Melissa Schettini, Thomas Tucker, Linda Tuch, and Lavelle (Yvette) Winfield served as our extremely able research assistants. David Howell of the Princeton University Department of East Asian Studies and Jin Heum Park kindly helped to determine the gender of Japanese and Korean names. We thank them all. We are grateful to our colleagues David Card, Anne Case, David Cutler, Angus Deaton, Hank Farber, Larry Katz, Alan Krueger, David Lee, and Aaron Yelowitz for helpful conversations, and to seminar participants at the School of Industrial and Labor Relations at Cornell University, University of Illinois at Champaign-Urbana, Princeton University, University of Toronto, Harvard University, and Vanderbilt University. We also thank two anonymous referees for comments that have made this a better paper. Any remaining errors are ours. Unfortunately the data used in this article are confidential and may not be made available to other researchers.

[1] An extensive literature exists on occupational segregation by sex and the possible reasons for the large differences in occupations between men and women today and in the past. The debate is ongoing. On the one hand are those who believe that discrimination, either individual or societal in nature, is the driving force, and on the other hand are those who claim the evidence shows women and men sort among occupations on the basis of different tastes for work characteristics. In the former category see Paula England (1982) and England et al. (1988); in the latter group see Solomon W. Polachek (1979) and Randall K. Filer (1989). It should be noted that many other studies (e.g., Ian Ayres and Joel Waldfogel, 1994) have also attempted to measure discrimination in atypical ways.

group of teachers. In an attempt to overcome this seeming bias in the hiring of musicians, most major U.S. orchestras changed their audition policies in the 1970's and 1980's making them more open and routinized. Openings became widely advertised in the union papers, and many positions attracted more than 100 applicants where fewer than 20 would have been considered before. Audition committees were restructured to consist of members of the orchestra, not just the conductor and section principal. The audition procedure became democratized at a time when many other institutions in America did as well.

But democratization did not guarantee impartiality, because favorites could still be identified by sight and through resumes. Another set of procedures was adopted to ensure, or at least give the impression of, impartiality. These procedures involve hiding the identity of the player from the jury. Although they take several forms, we use the terms "blind" and "screen" to describe the group.[2] The question we pose is whether the hiring process became more impartial through the use of blind auditions. Because we are able to identify sex, but no other characteristics for a large sample, we focus on the impact of the screen on the employment of women.[3]

Screens were not adopted by all orchestras at once. Among the major orchestras, one still does not have any blind round to their audition procedure (Cleveland) and one adopted the screen in 1952 for the preliminary round (Boston Symphony Orchestra), decades before the others. Most other orchestras shifted to blind preliminaries from the early 1970's to the late 1980's. The variation in screen adoption at various rounds in the audition process allows us to assess its use as a treatment.[4]

The change in audition procedures with the adoption of the screen allows us to test whether bias exists in its absence. In both our study and studies using audits, the issue is whether sex (or race or ethnicity), apart from objective criteria (e.g., the sound of a musical performance, the content of a resume), is considered in the hiring process. Why sex might make a difference is another matter.

Our data come from two sources: rosters and audition records. Rosters are simply lists of orchestra personnel, together with instrument and position (e.g., principal), found in orchestra programs. The audition records are the actual accounts of the hiring process kept by the personnel manager of the orchestra. Both are described in more detail below.

The audition records we have collected form an uncommon data set. Our sample includes who was advanced and hired from an initial group of contestants and also what happened to approximately two-thirds of the individuals in our data set who competed in other auditions in the sample. There are, to be certain, various data sets containing information on applicant pools and hiring practices (see, e.g., Harry Holzer and David Neumark, 1996). But our data set is unique because it has the complete applicant pool for each of the auditions and links individuals across auditions. Most important for our study is that audition procedures differed across orchestras in known ways and that the majority of the orchestras in our sample changed audition procedure during the period of study.[5]

We find, using our audition sample in an individual fixed-effects framework, that the screen increases the probability a woman will be advanced out of a preliminary round when there is no semifinal round. The screen also greatly enhances the likelihood a female contestant will be the winner in a final round. Using both the roster and auditions samples, and reasonable assumptions, the switch to blind auditions can explain about one-third of the increase in the proportion female among new hires (whereas another one-third is the result of the increased pool of female candidates). Estimates based on the roster sample indicate that blind auditions may account for 25 percent of the increase in the percentage of orchestra musicians who are female.

[2] For an article about the blind audition process see *The Economist* (1996).

[3] The screen may also have opened opportunities for individuals from less-well-known orchestras, those trained outside mainstream institutions, and those from minority groups.

[4] The blind audition procedures bear some resemblance to "double-blind" refereeing in academic journals. See Rebecca Blank (1991) for an assessment of the treatment effect of such refereeing in the *American Economic Review*.

[5] This statement is true for the roster sample. There are only a few orchestras that changed audition procedures during the years of our audition data.

I. Sex Composition of Orchestras

Symphony orchestras consist of about 100 musicians and, although the number has varied between 90 to 105, it is rarely lower or higher. The positions, moreover, are nearly identical between orchestras and over time. As opposed to firms, symphony orchestras do not vary much in size and have virtually identical numbers and types of jobs. Thus we can easily look at the proportion women in an orchestra without being concerned about changes in the composition of occupations and the number of workers. An increase in the number of women from, say, 1 to 10, cannot arise because the number of harpists (a female-dominated instrument), has greatly expanded. It must be because the proportion female within many groups has increased.

Among the five highest-ranked orchestras in the nation (known as the "Big Five")—the Boston Symphony Orchestra (BSO), the Chicago Symphony Orchestra, the Cleveland Symphony Orchestra, the New York Philharmonic (NYPhil), and the Philadelphia Orchestra—none contained more than 12 percent women until about 1980.[6] As can be seen in Figure 1A, each of the five lines (giving the proportion female) greatly increases after some point. For the NYPhil, the line steeply ascends in the early 1970's. For the BSO, the turning point appears to be a bit earlier. The percentage female in the NYPhil is currently 35 percent, the highest among all 11 orchestras in our sample after being the lowest (generally at zero) for decades. Thus the increase of women in the nation's finest orchestras has been extraordinary. The increase is even more remarkable because, as we discuss below, turnover in these orchestras is exceedingly low. The proportion of new players who were women must have been, and indeed was, exceedingly high.

Similar trends can be discerned for four other orchestras—the Los Angeles Symphony Orchestra (LA), the San Francisco Philharmonic (SF), the Detroit Symphony Orchestra, and the Pittsburgh Symphony Orchestra (PSO)—given in Figure 1B.[7] The upward trend in the proportion female is also obvious in Figure 1B, although initial levels are higher than in Figure 1A. There is somewhat more choppiness to the graph, particularly during the 1940's. Although we have tried to eliminate all substitute, temporary, and guest musicians, especially during World War II and the Korean War, this was not always possible.

The only way to increase the proportion women is to hire more female musicians and turnover during most periods was low. The number of new hires is graphed in Figure 2 for five orchestras. Because "new hires" is a volatile construct, we use a centered five-year moving average. In most years after the late 1950's, the top-ranked orchestras in the group (Chicago and NYPhil) hired about four musicians a year, whereas the other three hired about six. Prior to 1960 the numbers are extremely high for LA and the PSO, because, it has been claimed, their music directors exercised their power to terminate, at will, the employment of musicians. Also of interest is that the number of new hires trends down, even excluding years prior to 1960. The important points to take from Figure 2 are that the number of new hires was small after 1960 and that it declined over time.

The proportion female among the new hires must have been sizable to increase the proportion female in the orchestras. Figure 3 shows the trend in the share of women among new hires for four of the "Big Five" (Figure 3A) and four other orchestras (Figure 3B).[8] In both groups the female share of new hires rose over time, at a somewhat steeper rate for the more prestigious orchestras. Since the early 1980's the share female among new hires has been about 35 percent for the BSO and Chicago, and about 50 percent for the NYPhil, whereas before 1970 less than 10 percent of new hires were women.[9]

Even though the fraction of new hires who are female rises at somewhat different times

[6] The data referred to, and used in Figures 1 to 3, are from orchestral rosters, described in more detail below.

[7] Our roster sample also includes the Metropolitan Opera Orchestra and the St. Louis Symphony.

[8] A centered five-year moving average is also used for this variable.

[9] In virtually all cases the share of women among new hires has decreased in the 1990's.

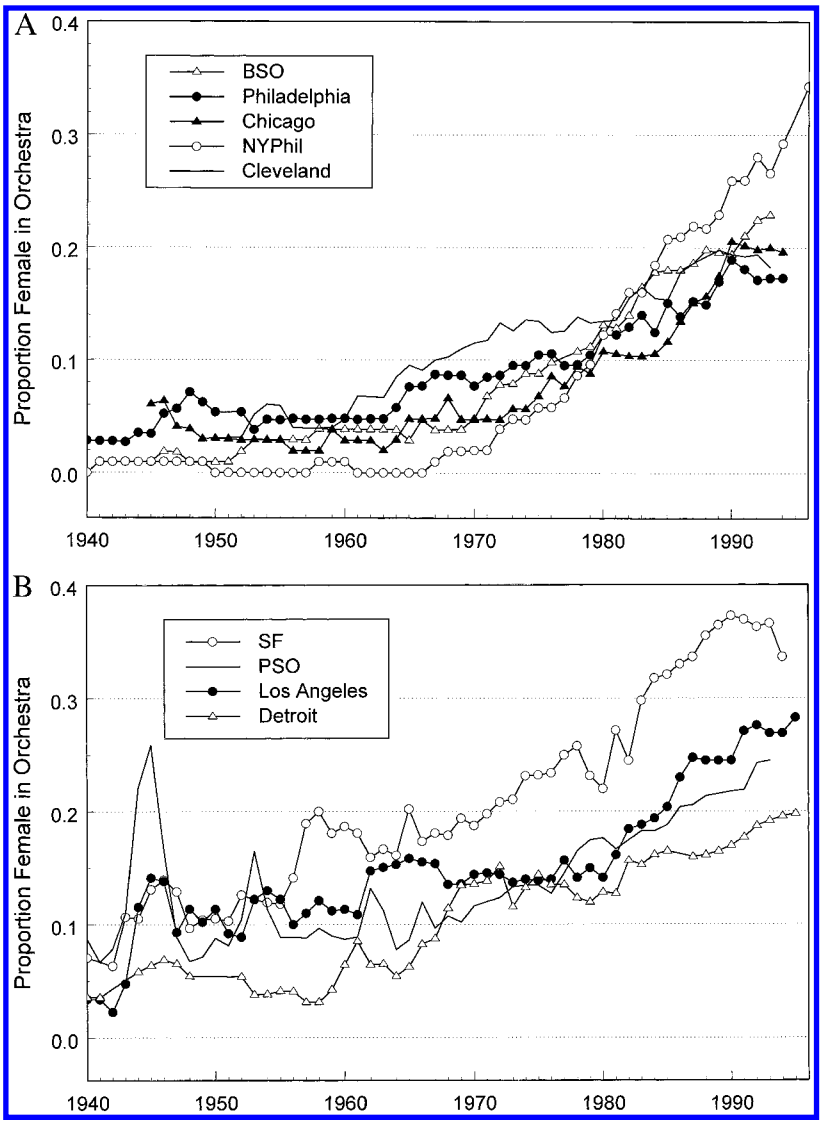

FIGURE 1. PROPORTION FEMALE IN NINE ORCHESTRAS, 1940 TO 1990's
A: THE "BIG FIVE"; B: FOUR OTHERS

Source: Roster sample. See text.

across the orchestras, there is a discernible increase for the group as a whole in the late 1970's to early 1980's, a time when the labor force participation of women increased generally and when their participation in various professions greatly expanded. The question, therefore, is whether the screen mattered in a direct manner or whether the increase was the result of a host of other factors, including the appearance of impartiality or an increased pool of female contestants coming out of music schools. Because the majority of new hires are in their late twenties and early thirties, the question is whether the most selective music schools were producing considerably more female students in the early 1970's. We currently have information by instrument for only the Juilliard School of Music. With the exception of the brass section, the data, given in Figure 4, do not reveal

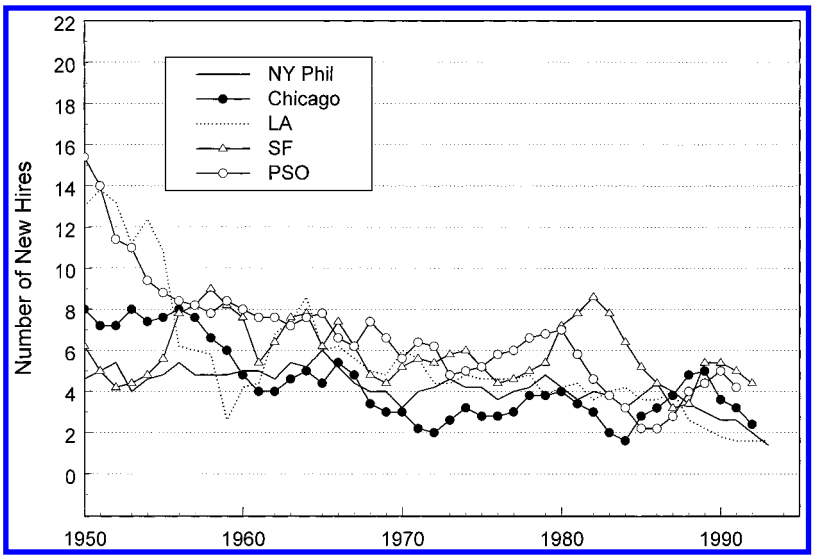

FIGURE 2. NUMBER OF NEW HIRES IN FIVE ORCHESTRAS, 1950 TO 1990's

Source: Roster sample. See text.

Notes: A five-year centered moving average is used. New hires are musicians who were not with the orchestra the previous year, who remain for at least one additional year, and who were not substitute musicians in the current year.

any sharp breaks in the fraction of all graduates who are female.[10] Thus, it is not immediately obvious that an expansion in the supply of qualified female musicians explains the marked increase in female symphony orchestra members; it could, therefore, be because of changes in the hiring procedures of orchestras.

But why would changes in audition procedures alter the sex mix of those hired? Many of the most renowned conductors have, at one time or another, asserted that female musicians are not the equal of male musicians. Claims abound in the world of music that "women have smaller techniques than men," "are more temperamental and more likely to demand special attention or treatment," and that "the more women [in an orchestra], the poorer the sound."[11] Zubin Mehta, conductor of the Los Angeles Symphony from 1964 to 1978 and of the New York Philharmonic from 1978 to 1990, is credited with saying, "I just don't think women should be in an orchestra."[12] Many European orchestras had, and some continue to have, stated policies not to hire women.[13] The Vienna Philharmonic has only recently admitted its first female member (a harpist). Female musicians, it can be convincingly argued, have historically faced considerable discrimination.[14] Thus a blind hiring procedure, such as a screen that conceals the identity of the musician auditioning, could eliminate

[10] We also have data on the sex composition of the graduates of the University of Michigan School of Music and Indiana University, but not by instrument. In the Michigan data, both for those receiving the Bachelor of Music (BM) degree and for those receiving the Master of Music (MM) degree, there is no change in the percentage female from 1972 to 1996. The Indiana University data, for both BM and MM degrees and excluding voice, piano, guitar, and early instruments, show an increase in the fraction female from 1975 to 1996. The ratio of females to males was 0.9 in 1975 but 1.2 in 1996.

[11] Seltzer (1989), p. 215.

[12] Seltzer (1989), p. 215. According to Seltzer, the fact that new hires at the NYPhil were about 45 percent female during Mehta's tenure as conductor suggests that Mehta's views may have changed.

[13] In comparison with the United Kingdom and the two Germanys, the United States in 1990 had the highest percentage female among its regional symphony orchestras and was a close second to the United Kingdom in the major orchestra category (Jutta J. Allmendinger et al., 1996).

[14] In addition, an African-American cellist (Earl Madison) brought a civil suit against the NYPhil in 1968 alleging that their audition procedures were discriminatory because they did not use a screen. The orchestra was found not guilty of discriminating in hiring permanent musicians, but it was found to discriminate in hiring substitutes.

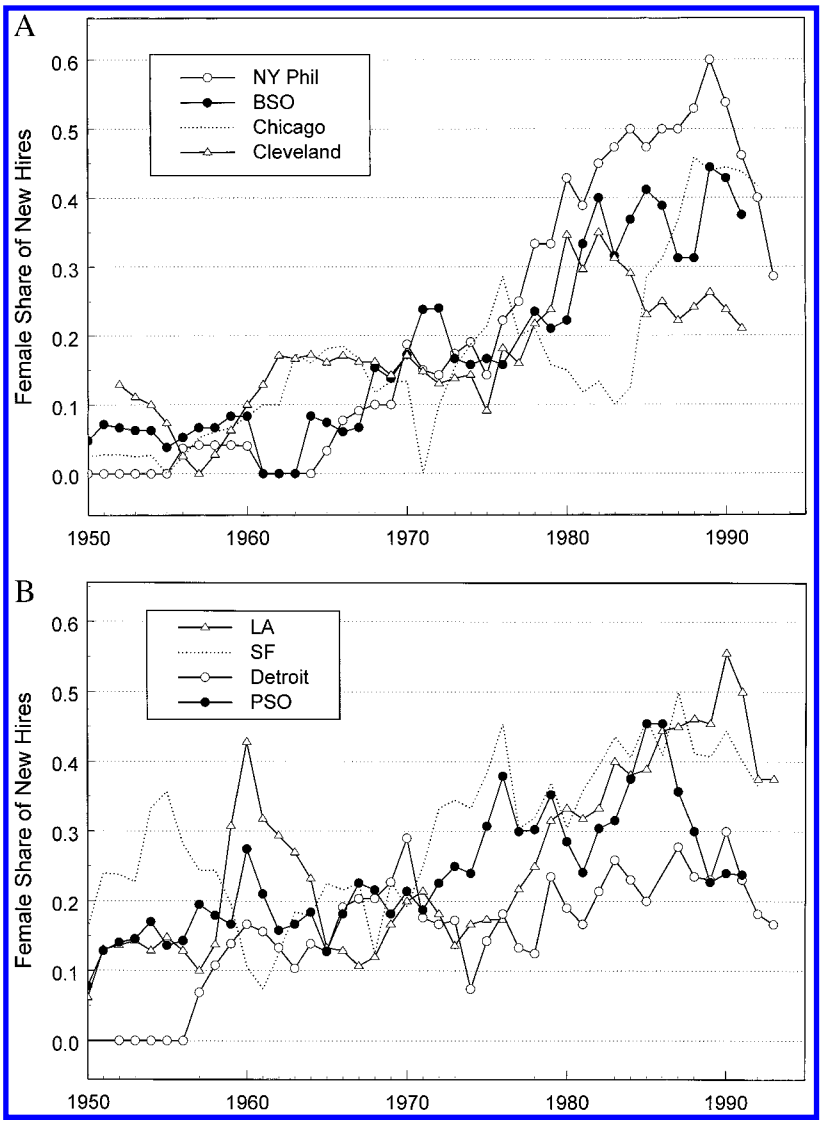

FIGURE 3. FEMALE SHARE OF NEW HIRES IN EIGHT ORCHESTRAS, 1950 TO 1990's
A: FOUR OF THE "BIG FIVE"; B: FOUR OTHERS

Source: Roster sample. See text.

Notes: A five-year centered moving average is used. New hires are musicians who were not with the orchestra the previous year, who remain for at least one additional year, and who were not substitute musicians in the current year.

the possibility of discrimination and increase the number of women in orchestras.

II. Orchestral Auditions

To understand the impact of the democratization of the audition procedure and the screen, we must first explain how orchestra auditions are now conducted. After determining that an audition must be held to fill an opening, the orchestra advertises that it will hold an audition. Each audition attracts musicians from across the country and, often,

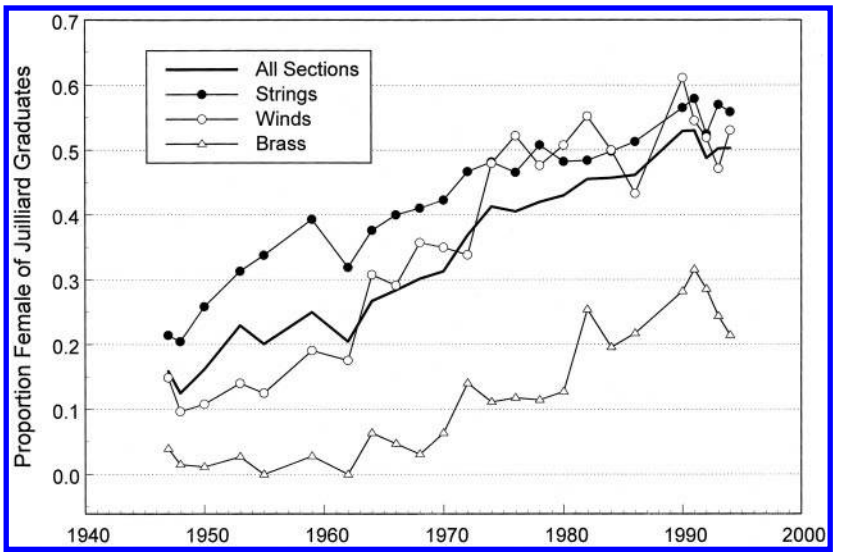

FIGURE 4. PROPORTION FEMALE OF JUILLIARD GRADUATES, TOTAL AND BY SECTION: 1947 TO 1995

Source: Juilliard Music School files.

from around the world.[15] Musicians interested in auditioning are required to submit a resume and often a tape of compulsory music (recorded according to specific guidelines) to be judged by members of the orchestra. In some orchestras this prescreening is dispositive; in others the musician has the right to audition live in the preliminary round, even if the audition committee rejects the candidate on the basis of the tape.[16] All candidates are given, in advance, most of the music they are expected to perform at the live audition.

Live auditions today generally consist of three rounds: preliminary, semifinal, and final. But there is considerable variation. Although all orchestras now have a preliminary round, some have two final rounds and in many there was no semifinal round until the 1980's. The preliminary is generally considered a screening round to eliminate unqualified candidates. As a result, the committee is free to advance as many, or as few, as they wish. Candidates advanced from the semifinal round are generally considered "acceptable for hire" by the audition committee (which does not include the music director, a.k.a. conductor, until the finals). Again, this means that the committee can advance as many as it wishes. The final round generally results in a hire, but sometimes does not.[17]

In blind auditions (or audition rounds) a screen is used to hide the identity of the player from the committee.[18] The screens we have seen are either large pieces of heavy (but sound-porous) cloth, sometimes suspended from the ceiling of the symphony hall, or look like large room dividers. Some orchestras also roll out a carpet leading to center stage to muffle footsteps that could betray the sex of the candidate.[19] Each candidate for a blind audition is given a number, and the jury rates the candidate's

[15] Orchestral auditions, particularly for the nation's most prestigious orchestras, are national if not international, in scope. Many contestants, the vast majority of whom receive no travel reimbursement, travel long distances to audition. The auditions span the fewest number of days possible to minimize hotel charges.

[16] The tape, in this case, provides information to the candidate of his or her likelihood of success, sparing the musician a potentially large travel expense.

[17] There is one exception to this general rule. In rare cases when the committee cannot decide between two or three candidates, each is invited to play with the orchestra before the final decision is made.

[18] It may also serve to hide the identity of the committee from the player, although that is not its main function. We use the terms "blind" and "screen" interchangeably.

[19] Or, if a carpet is not placed on the stage, the personnel manager may ask a woman to take off her shoes and he provides the compensating footsteps.

performance next to the number on a sheet of paper. Only the personnel manager knows the mapping from number to name and from name to other personal information.[20] The names of the candidates are not revealed to the juries until after the last blind round.

Almost all preliminary rounds are now blind. The semifinal round, added as the number of applicants grew, may be blind. Finals are rarely blind and almost always involve the attendance and input of the music director.[21] Although the music director still wields considerable power, the self-governance that swept orchestras in the 1970's has served to contain the conductor's authoritarianism. The music director can ignore the audition committee's advice, but does so at greater peril. Once an applicant is chosen to be a member of an orchestra, lifetime tenure is awarded after a brief probationary period. The basis for termination is limited and rarely used. The positions we are analyzing are choice jobs in the musical world. In 1995 the *minimum starting* base salary for musicians at the BSO was $1,400 per week (for a 52-week year), not including recording contracts, soloist fees, overtime and extra service payments, bonuses, and per diem payments for tours and Tanglewood.[22]

Are blind auditions truly blind, or can a trained, accomplished musician identify contestants solely from differences in playing style, just as academics can often identify authors of double-blind papers they get to referee? Unlike double-blind refereeing, for which one sees an entire paper with its distinctive writing style, methodology, sources, and citations, the candidates play only predetermined and brief excerpts from the orchestral repertoire. Each candidate typically has just 5 to 10 minutes to play for the audition committee, particularly in the early rounds. There is little or no room for individuality to be expressed and not much time for it to be detected.[23] Even when an individual musician is known in advance to be auditioning, jury members often cannot identify that individual. Only the rare, well-known candidate, with an unusually distinctive musical style could conceivably be correctly identified.

The many musicians and personnel managers with whom we have spoken uniformly deny that identification is possible for the vast majority of contestants. They also observe that, although it is tempting to guess the identity of a contestant, particularly in the later rounds, audition committee members, more often than not, find they are wrong. To base a hiring decision on speculation would not be in the best interests of the orchestra. Further, although an individual committee member may believe that he or she knows the identity of a player, it would be rare for the entire committee to be secure in such knowledge. Thus, even if one committee member's vote is swayed by such a belief, the committee's vote must correspond to the consensus view of the player's musical ability for it to determine the outcome. Thus, auditions held with a screen, apart from very few exceptions, are truly blind.

The audition procedures of the 11 orchestras in the roster sample are summarized in Table 1.[24] Although audition procedures are now part of union contracts, that was not the case in the more distant past and the procedures were not apparently recorded in any surviving documents. We gathered information on these procedures from various sources, including union contracts, interviews with personnel managers, archival documents on auditions, and a mail survey we conducted of orchestral musicians concerning the proce-

[20] The personnel manager is generally a musician who played with the orchestra for some time and knows the players and the conductor well. The duties involve managing the day-to-day work of the orchestra, getting substitute musicians, making travel plans, and arranging the hiring of new musicians.

[21] It is almost always the case that if an orchestra in, say, the spring of 1986 holds a blind preliminary round for a position, it will have all its candidates audition blind in that round and in all other preliminary rounds during that season, should there be any. That is, there is generally no discretion on the part of the jury (and certainly not on the part of the contestant) in terms of the audition procedure, particularly once an audition is underway.

[22] Most of the orchestra contracts in the group we have examined have similar base salaries. Union contracts list only the minimum or base starting salary and minimum increments for seniority. We do not know how many musicians have individually negotiated rates above the stated minimum amounts.

[23] Also, there is generally not a standing audition committee that might become familiar with the musicians who audition frequently.

[24] We identify the orchestras by letter, rather than by name, to preserve confidentiality of the audition sample.

TABLE 1—ORCHESTRA AUDITION PROCEDURE SUMMARY TABLE

Orchestra	Preliminaries	Semifinals	Finals
A	Blind since 1973	Blind (varies) since 1973	Not blind
B	Blind since at least 1967	Use of screen varies	Blind 1967–1969; since winter 1994
C	Blind since at least 1979 (definitely after 1972)	Not blind: 1991–present Blind: 1984–1987	Not blind
D	Blind since 1986	Blind since 1986; varies until 1993	1st part blind since 1993; 2nd part not blind
E	Use of screen varies until 1981	Use of screen varies	Not blind
F	Blind since at least 1972	Blind since at least 1972	Blind since at least 1972
G	Blind since 1986	Use of screen varies	Not blind
H	Blind since 1970	Not blind	Not blind
I	Blind since 1979	Blind since 1979	Blind since fall 1983
J	Blind since 1952	Blind since 1952	Not blind
K	Not blind	Not blind	Not blind

Notes: The 11 orchestras (A through K) are those in the roster sample described in the text. A subset of eight form the audition sample (also described in the text). All orchestras in the sample are major big-city U.S. symphony orchestras and include the "Big Five."

Sources: Orchestra union contracts (from orchestra personnel managers and libraries), personal conversations with orchestra personnel managers, and our mail survey of current orchestra members who were hired during the probable period of screen adoption.

dures employed during the audition that won them their current position.

An obvious question to ask is whether the adoption of the screen is endogenous. Of particular concern is that more meritocratic orchestras adopted blind auditions earlier, producing the spurious result that the screen increased the likelihood that women were hired.[25] We estimate a probit model of screen adoption by year, conditional on an orchestra's not previously having adopted the screen (an orchestra exits the exercise once it adopts the screen). Two time-varying covariates are included to assess commonly held notions about screen adoption: the proportion female (lagged) in the orchestra, and a measure of tenure (lagged) of then-current orchestra members. Tenure is included because personnel managers maintain the screen was advocated more by younger players.

As the proportion female in an orchestra increases, so does the likelihood of screen adoption in the preliminary round, as can be seen in columns (1) and (2) in Table 2, although the effects are very small and far from statistically significant.[26] We estimate a similar effect when we assess the role of female presence on the adoption of blind finals [see column (3)]. The impact of current tenure, measured by the proportion with less than six years with the orchestra, is—contrary to general belief—negative and the results do not change controlling for whether the orchestra is one of the "Big Five."[27] In all, it appears that orchestra sex composition had little influence on screen adoption, although the stability of the personnel may have increased its likelihood.[28]

[25] Note, however, it is unlikely that the orchestras that sought to hire more women chose to adopt the screen earlier since the best way to increase the number of women in the orchestra is to have not-blind auditions (so that one could be sure to hire more women).

[26] An increase in the proportion female from 0 to 0.35, the largest for any of the orchestras (see Figure 1), would enhance the likelihood of adopting the screen in the preliminary round by a mere 0.0021 percentage points.

[27] Our measure of tenure begins at the first date for which we have rosters, but not earlier than 1947. Tenure then cumulates for each member until the individual exits the orchestra. Because tenure will increase for all orchestras with time, we use the proportion of all members with fewer than six years of tenure.

[28] A change in conductor could also have led to a change in the audition policy, but we find no supporting evidence. For example, current players contend that Charles Munch had complete authority in hiring at the BSO before 1952. The BSO adopted the screen in 1952, but Munch was

TABLE 2—ESTIMATED PROBIT MODELS FOR THE USE OF A SCREEN

	Preliminaries blind		Finals blind
	(1)	(2)	(3)
(Proportion female)$_{t-1}$	2.744	3.120	0.490
	(3.265)	(3.271)	(1.163)
	[0.006]	[0.004]	[0.011]
(Proportion of orchestra	−26.46	−28.13	−9.467
personnel with <6	(7.314)	(8.459)	(2.787)
years tenure)$_{t-1}$	[−0.058]	[−0.039]	[−0.207]
"Big Five" orchestra		0.367	
		(0.452)	
		[0.001]	
pseudo R^2	0.178	0.193	0.050
Number of observations	294	294	434

Notes: The dependent variable is 1 if the orchestra adopts a screen, 0 otherwise. Huber standard errors (with orchestra random effects) are in parentheses. All specifications include a constant. Changes in probabilities are in brackets. "Proportion female" refers to the entire orchestra. "Tenure" refers to years of employment in the current orchestra. "Big Five" includes Boston, Chicago, Cleveland, New York Philharmonic, and Philadelphia. The data begin in 1947 and an orchestra exits the exercise once it adopts the screen. The unit of observation is an orchestra-year.
Source: Eleven-orchestra roster sample. See text.

III. The Role of Blind Auditions on the Audition and Hiring Process

A. *Data and Methods*

Audition Records.—We use the actual audition records of eight major symphony orchestras obtained from orchestra personnel managers and the orchestra archives. The records are highly confidential and occasionally contain remarks (including those of the conductor) about musicians currently with the orchestra. To preserve the full confidentiality of the records, we have not revealed the names of the orchestras in our sample.

Although availability differs, taken together we obtained information on auditions dating from the late 1950's through 1995. Typically, the records are lists of the names of individuals who attended the auditions, with notation near the names of those advanced to the next round. For the preliminary round, this would indicate advancement to either the semifinal or final round. Another list would contain the names of the semifinalists or finalists with an indication of who won the audition.[29] From these records, we recorded the instrument and position (e.g., section, principal, substitute) for which the audition was held. We also know whether the individual had an "automatic" placement in a semifinal or final round. Automatic placement occurs when a musician is already known to be above some quality cutoff and is invited to compete in a semifinal or final round.[30] We also recorded whether the individual was advanced to the next round of the current audition.

We rely on the first name of the musicians to determine sex. For most names establishing sex was straightforward.[31] Sexing the Japanese and Korean names was equally straightforward, at least for our Japanese and Korean consultants. For more difficult cases, we checked the names in three baby books (Connie Lockhard Ellefson, 1990; Alfred J. Kolatch, 1990; Bruce Lansky, 1995). If the name was listed as male- or female-only, we considered the sex known. The gender-neutral names (e.g., Chris, Leslie, and Pat) and some Chinese names (for which sex is indeterminate in the absence of Chinese characters) remained ambiguous. Using these procedures, we were able to determine the sex of 96 percent of our audition sample.[32] We later assess the impact that sex misclassification may have on our results.

In constructing our analysis sample, we exclude incomplete auditions, those in which there were no women (or only women) competing, rounds from which no one was advanced, and the second final round, if one exists, for which

conductor from 1949 to 1962. Our inability to explain the timing of screen adoption may result from our lack of intimate knowledge of the musical world, although it is also difficult to explain blind refereeing policy among economics journals (see the list in Blank, 1991).

[29] In rare cases, we have additional information on the finalists, such as resumes.

[30] The person will be known to be above a quality cutoff either because the individual is a current member of a comparable orchestra or because the person was a semifinalist or finalist in a previous audition.

[31] For 13 percent of the contestants, sex was confirmed by personnel managers, resumes, or audition summary sheets.

[32] Most of the remainder were sexed using census data by assigning to them the dominant sex of individuals with their first name.

TABLE 3—DESCRIPTIVE STATISTICS ABOUT AUDITIONS, BY YEAR AND ROUND OF AUDITION

Year	Number of auditions	Proportion female	Number of musicians	Number of auditions	Proportion female	Number of musicians	Number of auditions	Proportion female
			Completely blind auditions			Not completely blind auditions		
All	254	0.367	43.4	60	0.393	38.1	194	0.359
		(0.013)	(3.13)		(0.029)	(1.74)		(0.015)
Pre-1970	10	0.187				16.3	10	0.187
		(0.042)				(2.27)		(0.042)
1970–1979	69	0.329				31.4	69	0.329
		(0.026)				(2.10)		(0.026)
1980–1989	102	0.394	42.5	33	0.375	39.6	69	0.403
		(0.019)	(4.29)		(0.034)	(2.73)		(0.022)
1990+	73	0.390	44.6	27	0.415	50.6	46	0.375
		(0.027)	(4.64)		(0.049)	(4.52)		(0.033)
Round				Blind rounds			Not-blind rounds	
Preliminaries, without semifinals	170	0.357	34.3	125	0.367	24.7	45	0.327
		(0.015)	(1.87)		(0.017)	(2.33)		(0.029)
Preliminaries, with semifinals	137	0.396	45.5	134	0.395	49.3	3	0.425
		(0.019)	(2.54)		(0.019)	(17.0)		(0.205)
Semifinals	114	0.415	12.3	89	0.404	10.4	25	0.455
		(0.019)	(0.649)		(0.022)	(1.21)		(0.043)
Finals	167	0.430	4.93	28	0.472	7.12	130	0.422
		(0.016)	(0.448)		(0.040)	(0.310)		(0.017)

Notes: The unit of observation for the top portion is the audition, whereas it is the round for the bottom portion (e.g., proportion female in the top portion of the table is averaged across the auditions). Standard errors are in parentheses.
Source: Eight-orchestra audition sample. See text.

the candidates played with the orchestra.[33] In addition, we generally consider each round of the audition separately. These sample restrictions exclude 294 rounds (199 contained no women) and 1,539 individuals. Our final analysis sample has 7,065 individuals and 588 audition rounds (from 309 separate auditions) resulting in 14,121 person-rounds and an average of 2.0 rounds per musician.[34]

As can be seen in the bottom portion of Table 3, 259, or 84 percent, of our 307 preliminary rounds were blind, 78 percent of the 114 semifinals were blind, but just 17 percent of the 167 final rounds were blind. Most of our audition sample is for the period after 1970. The blind preliminaries contained 40 candidates on average, whereas those without the screen had 26. Women were about 37 percent of all preliminary candidates but 43 percent of finalists, and the difference holds for both the blind and not-blind auditions. The percentage female among all candidates increased over time, from 33 percent in the 1970 to 1979 period to 39 percent in the post-1990 years (see upper portion).

Roster Data.—Our second source of information comes from the final results of the audition process, the orchestra personnel rosters. We collected these data from the personnel page of concert programs, one each year for eleven major symphony orchestras. These records are in the public domain and thus we have used the orchestra names in the graphs containing those data alone. As opposed to the auditionees, we were able to confirm the sex of the players with the orchestra personnel managers and

[33] Although the results are unaffected, harp auditions are excluded because it has typically been a female-dominated instrument.

[34] See Table A1 for descriptive statistics.

archivists. We considered a musician to be new to the orchestra in question if he or she had not previously been a regular member of that orchestra (i.e., we did not count returning members as new). We excluded, when possible, temporary and substitute musicians, as well as harpists and pianists. Our final sample for 1970 to 1996 has 1,128 new orchestra members (see Table A2).

Econometric Framework.—We take advantage of the variation that exists across orchestras, time, and audition round to identify the effect of the screens on the likelihood that a female is advanced from one round to the next and ultimately hired. The probability that individual i is advanced (or hired) from an audition at orchestra j, in year t, from round r, is a function of the individual's sex (F), whether a screen is used (B), and other individual (X) and orchestral (Z) factors, that is:

$$(1) \qquad P_{ijtr} = f(X_{it}, F_i, B_{jtr}, Z_{jtr}).$$

The screen, it will be recalled from Table 1, varies across orchestra, time, and audition round. Orchestras adopted the screen in different years. Some used the screen in the preliminary round only, whereas others used the screen for the entire audition process. We use this variation to estimate a differences-in-differences strategy. In linear form, we write

$$(2) \quad P_{ijtr} = \alpha + \beta F_i + \gamma B_{jtr} + \delta(F_i \times B_{jtr}) + X_{it}\theta_1 + Z_{jtr}\theta_2 + \varepsilon_{ijtr}.$$

The coefficient on B_{jtr}, γ, identified from the men who audition with a screen, controls for whether all individuals are more or less likely to be advanced from a blind than from a not-blind audition. Thus the parameter of interest is that on the interaction between F_i and B_{jtr}, δ, which measures the change in the probability that a woman will be advanced if a screen is used, relative to her auditioning without a screen (after accounting for other blind audition effects). We also test whether the use of the screen eliminates sex differences in the likelihood an individual is advanced from one round to the next. Because no restrictions exist on the number of individuals advanced from the preliminary and semifinal rounds, there is no zero-sum game between men and women for these rounds.

B. *The Effect of the Screen on the Likelihood of Being Advanced*

Tabulations and Regression Results With and Without Individual Fixed Effects.—The raw data in Tables 4 and 5 can reveal the impact on women of changes in the audition process and provide an important introduction to the data. We demonstrate that in the absence of a variable for orchestral "ability," women fare *less* well in blind auditions than otherwise. But if the orchestral "ability" of the candidate is held fixed, the screen provides an unambiguous and substantial *benefit* for women in almost all audition rounds.

Table 4 gives the success rate by sex, round of audition, and over time. We define "relative female success" as the proportion of women advanced (or hired) minus the proportion of men advanced (or hired). The relative success of female candidates appears worse for blind than for not-blind auditions and this finding also holds for each round of the audition process. One interpretation of this result is that the adoption of the screen lowered the average quality of female auditionees in the blind auditions. Only if we can hold quality constant can we identify the true impact of the screen.

Because we have the names of the candidates, we are able to link their success in one audition to that in another. (In our sample, 24 percent of the individuals competed in more than one audition.) In Table 5 we report audition success statistics, by round and overall, for musicians who appear more than once in our sample and for whom at least one audition (or round) was blind and one was not blind. The evidence tells a very different story from that in Table 4, and taken together they suggest that blind auditions expanded the pool of female applicants to include more who were less qualified. When we limit the sample to those who auditioned both with and without a screen, the success rate for women competing in blind auditions is almost always higher than in those that were not blind.

TABLE 4—AVERAGE SUCCESS AT AUDITIONS BY SEX, YEAR, AND ROUND OF AUDITION

	Relative female success		
Year	All auditions	Completely blind auditions	Not completely blind auditions
All	−0.001	−0.022	0.006
	(0.008)	(0.012)	(0.010)
Pre-1970	0.053		0.053
	(0.115)		(0.115)
1970–1979	0.001		0.001
	(0.021)		(0.021)
1980–1989	−0.006	−0.039	0.010
	(0.009)	(0.016)	(0.009)
1990+	−0.003	−0.001	−0.003
	(0.010)	(0.017)	(0.013)
Round	All rounds	Blind rounds	Not-blind rounds
Preliminaries, without semifinals	−0.032	−0.048	0.012
	(0.019)	(0.021)	(0.040)
Preliminaries, with semifinals	−0.048	−0.052	0.116
	(0.016)	(0.016)	(0.228)
Semifinals	−0.030	−0.059	0.071
	(0.038)	(0.044)	(0.080)
Finals	0.009	−0.028	0.016
	(0.036)	(0.102)	(0.038)

Notes: For the top part of the table "success" is a "hire," whereas for the bottom portion "success" is advancement from one stage of an audition to the next. The unit of observation for the top portion is the audition, whereas it is the round for the bottom portion (e.g., relative female success in the top portion of the table is averaged across the auditions). Standard errors are in parentheses. "Relative female success" is the proportion of women advanced (or hired) minus the proportion of men advanced (or hired). By hired, we mean those who were advanced from the final round out of the entire audition.
Source: Eight-orchestra audition sample. See text.

Take the preliminary round with no semifinals, for example, in Table 5. In the blind auditions 28.6 percent of the women are advanced, as are 20.2 percent of the men. But in the not-blind column, just 19.3 percent of the women are advanced, although 22.5 percent of the men are. Even though a woman has a small advantage over a man when the screen is used (by 8.4 percentage points), her success rate, relative to that of a man, is increased by 11.6 percentage points above that in the not-blind regime. Note that because these are the *same* women, Table 5 suggests that a woman enhances her own success rate by 9.3 percentage points by entering a blind preliminary round. Not only do these differences suggest that women are helped by the screen, the differences are large relative to the average rate of success.[35]

Women's success is also enhanced by the screen in the finals and for the overall audition (termed "hired" in the table). For the finals, a woman's success rate is increased by 14.8 percentage points moving to blind auditions (23.5 − 8.7) and is enhanced by a hefty 28.1 percentage points above that of men. All success rates are very low for auditions as a whole, but the female success rate is 1.6 times higher (increasing from 0.017 to 0.027) for blind than for not-blind auditions. The only anomalous result in the table concerns the semifinals, to which we return later. We now show that these results stand up to the controls we can add, including the year of the audition and the instrument.[36]

[35] Because of the infrequency of position availability, it is unlikely there was much gaming by women (e.g., trying out only for blind auditions), although the change in the general environment of auditions could have altered the pool of contestants.

[36] We do not discuss the regression analog to Table 4, that is, the analysis without individual fixed effects, because we have firmly established that individual fixed effects matter. Table A3 shows the results of regressions

TABLE 5—AVERAGE SUCCESS AT AUDITIONS BY SEX AND STAGE OF AUDITION FOR THE SUBSET OF MUSICIANS WHO AUDITIONED BOTH BLIND AND NOT BLIND

	Blind		Not blind	
	Proportion advanced	Number of person-rounds	Proportion advanced	Number of person-rounds
	Preliminaries without semifinals			
Women	0.286	112	0.193	93
	(0.043)		(0.041)	
Men	0.202	247	0.225	187
	(0.026)		(0.031)	
	Preliminaries with semifinals			
Women	0.200	20	0.133	15
	(0.092)		(0.091)	
Men	0.083	12	0.000	8
	(0.083)		(0.000)	
	Semifinals			
Women	0.385	65	0.568	44
	(0.061)		(0.075)	
Men	0.368	68	0.295	44
	(0.059)		(0.069)	
	Finals			
Women	0.235	17	0.087	23
	(0.106)		(0.060)	
Men	0.000	12	0.133	15
	(0.000)		(0.091)	
	"Hired"			
Women	0.027	445	0.017	599
	(0.008)		(0.005)	
Men	0.026	816	0.027	1102
	(0.005)		(0.005)	

Notes: The unit of observation is a person-round. Standard errors are in parentheses. For the round in question, only musicians who auditioned more than once and who auditioned at least once behind a screen and at least once without a screen are included. "Hired" means those who were advanced from the final round out of the entire audition. Blind in the "hired" category means for all rounds. Not blind in the "hired" category means that at least one round was not blind. This difference in the definition of what constitutes a "blind" round or audition is one reason why the number of observations in the first four panels is less than the number of observations in the "hired" panel. The number of observations also differ because we exclude auditions or rounds in which no individual is advanced or in which there are only women or no women. Finally, unlike in subsequent tables, we exclude a few candidates for whom we could not determine or impute their sex. Note that the binding constraint for the preliminaries is the not-blind category, for which we have only one orchestra. The binding constraint in the "hired" category are the blind auditions, for which we have (at most) three orchestras. Musicians can appear more than once in either the blind or not-blind categories.
Source: Eight-orchestra audition sample. See text.

The results given in Table 6 are the regression analogs to the raw tabulations in Table 5.[37] Because the effect of the blind procedure could differ by the various rounds in the audition process, we divide audition rounds into the three main rounds (preliminary, semifinal, and final) and also separate the preliminaries into those that were followed by a semifinal

comparable to those in Table 6 but without individual fixed effects.

[37] In the (total) subsample of individuals auditioning both with and without a screen, all eight orchestras in our audition sample are represented, and seven of the orchestras changed audition policy during our sample time frame. The sample sizes in Table 6 are considerably larger than those in Table 5. The reason is that the regressions in Table 6 include *all* individuals whether or not they auditioned more than once, whereas Table 5 includes only those who auditioned at least twice, blind *and* not blind.

TABLE 6—LINEAR PROBABILITY ESTIMATES OF THE LIKELIHOOD OF BEING ADVANCED: WITH INDIVIDUAL FIXED EFFECTS

	Preliminaries							
	Without semifinals		With semifinals		Semifinals		Finals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Blind	−0.017	0.003	0.109	0.224	0.026	0.102	−0.154	−0.060
	(0.039)	(0.046)	(0.172)	(0.242)	(0.089)	(0.096)	(0.150)	(0.149)
Female × Blind	0.125	0.111	0.013	−0.025	−0.179	−0.235	0.308	0.331
	(0.068)	(0.067)	(0.215)	(0.251)	(0.126)	(0.133)	(0.196)	(0.181)
Number of auditions attended		−0.020		0.010		0.015		0.126
		(0.014)		(0.010)		(0.030)		(0.028)
Years since last audition		−0.005		−0.006		−0.005		0.016
		(0.007)		(0.005)		(0.013)		(0.015)
Automatic placement						−0.096		−0.069
						(0.064)		(0.073)
"Big Five" orchestra		−0.154		−0.059		0.006		−0.059
		(0.035)		(0.024)		(0.081)		(0.084)
Total number of auditioners in round (÷100)		−0.003		0.014		−0.371		−0.262
		(0.081)		(0.031)		(0.521)		(0.756)
Proportion female at the audition round		0.118		0.312		0.104		0.067
		(0.139)		(0.134)		(0.218)		(0.159)
Principal		−0.079		−0.078		−0.082		−0.185
		(0.037)		(0.019)		(0.066)		(0.076)
Substitute		0.165		0.123		0.167		0.079
		(0.081)		(0.093)		(0.183)		(0.217)
p-value of H_0: Blind + (Female × Blind) = 0	0.053	0.063	0.342	0.285	0.089	0.170	0.222	0.042
Year fixed effects?	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.748	0.775	0.687	0.697	0.774	0.794	0.811	0.878
Number of observations	5,395	5,395	6,239	6,239	1,360	1,360	1,127	1,127

Notes: The unit of observation is a person-round. The dependent variable is 1 if the individual is advanced to the next round and 0 if not. Standard errors are in parentheses. All specifications include individual fixed effects, an interaction for the sex being missing and a blind audition round, a dummy indicating if years since last audition is missing, and [in columns (3)–(8)] whether an automatic placement is missing.
Source: Eight-orchestra audition sample. See text.

round and those that were not. In the even-numbered columns we include year and instrument fixed effects, as well as individual and audition covariates. The individual correlates are whether the musician had an automatic placement in a semifinal or final round, years since the last audition in the sample, and the number of previous auditions in which we observe the musician to have competed. We also control for the total number of musicians in the round, the proportion female among contestants, and whether the audition is for a principal or substitute position.

Because 42 percent of the individuals in our sample competed in more than one round in our data set (24 percent of the musicians competed in more than one audition) and 6 percent competed both with and without a screen for a particular type of round (e.g., semifinal), we are able to use an individual fixed-effects strategy to control for contestant "ability" that does not change with time. In all columns of Table 6 we include individual fixed effects, in which case the identification is from individuals who auditioned both with and without a screen.[38] The

[38] There are 639 person-rounds comprised of individuals who auditioned at a preliminary round that was not followed by a semifinal round [columns (1) and (2) of Table 6], both with and without a screen; on average these individuals competed in 2.7 such preliminary rounds. There are 55 person-rounds comprised of individuals who auditioned at a preliminary round that was followed by a semifinal round [columns (3) and (4)], both with and without a screen; on average these individuals competed in 2.4 such preliminary rounds. There are 223 person-rounds comprised of individuals who auditioned at a semifinal [columns (5) and (6)],

effect of the screen here, therefore, is identified from differing audition procedures both within and across orchestras.[39] Note that we include a dummy variable for whether the orchestra is among the "Big Five," to control for the quality of the orchestra.

The coefficient of interest is the interaction between "Female" and "Blind." A positive coefficient would show that screened auditions enhance a woman's likelihood of advancement. Because screened auditions are more likely to take place in later years than auditions without screens, the interaction between "Female" and "Blind" might simply reflect the fact that female musicians get better over time. Note, however, that for this effect to bias the coefficient, female musicians would have to improve faster with time than male musicians. Nevertheless, we have also included (in the individual covariates) the number of previous auditions the musician attended in our sample, the number of years since the last audition in the sample, and whether the candidate was an automatic placement. The coefficient on "Blind" reveals whether blind auditions change the likelihood that all contestants are advanced.

As in the raw tabulations of Table 5, we find that the screen has a *positive* effect on the likelihood that a woman is advanced from the preliminary round (when there is no semifinal) and from the finals.[40] The effects, moreover, are statistically significant in both cases. The effect in the semifinal round, however, remains strongly negative.[41] In addition, the magnitudes of the effects in Table 6 are similar to those implied by the raw tabulations (Table 5). For preliminaries that are not preceded by a semifinal, the blind audition increases the likelihood that a woman will be selected by about 11 percentage points. For female musicians who made it to the final round, the individual fixed-effects regression result indicates that the screen increases the likelihood of their winning by about 33 percentage points.[42]

Assessing Potential Biases.—A concern with the preceding fixed-effects analysis is that, as noted earlier, female musicians who are improving over time are those who switch from not-blind to blind auditions and that the growth rate of their "ability" is faster than that of men. We attempted to address this potential bias by including several individual time-varying co-

both with and without a screen; on average these individuals participated in 2.8 semifinal rounds. Finally, there are 67 person-rounds comprised of individuals who auditioned at a final round [columns (7) and (8)], both with and without a screen; on average these individuals participated in 2.4 final rounds. It should be noted that the number of person-rounds off of which we are identified in Table 6 can also be found in Table 5, with one exception. There are 223 person-rounds comprised of individuals who auditioned at the semifinal, both with and without a screen, in Table 6 and only 221 in Table 5 because there are two individuals we could not sex. We include these individuals in the regressions in Table 6 and add a dummy variable indicating that the sex is missing.

[39] An analysis of variance (ANOVA) across the entire sample, that is pooling all rounds, indicates that 19 percent of the variation in the use of the screen is across orchestras. Looking by audition round reveals that 73 percent of the variation in preliminaries, 53 percent of the variation in semifinals, and 71 percent of the variation in finals is across orchestras. By contrast, in Table 7 (which includes a subset of the orchestras, see table notes), just 1 percent of the variation in the use of the screen is across orchestras.

[40] An exception occurs when preliminaries are followed by semifinals. There are, however, only three preliminary rounds that are not blind when there is also a semifinal round (see Table 3). Thus the coefficients in columns (3) and (4) of Table 6 are identified using very few separate audition rounds. We also note that when we estimate fixed-effects logit models we obtain results similar to those in columns (1) and (2) in Table 6 (and in Table 7). Because of the small samples with the identifying requirements of the fixed-effects logit, standard errors for the estimates in columns (3)–(8) of Table 6 could not be computed. Further, for the results without individual fixed effects, logits and linear probability models give qualitatively similar results.

[41] This result on the semifinals is robust across time, instrument, position, and orchestra. One interpretation is that it represents a form of affirmative action by the audition committees. Committees may hesitate to advance women from the preliminary round if they are not confident of the candidate's ability. On the other hand, semifinals are typically held the same day as are preliminaries and give the audition committee a second chance to hear a candidate before the finals. Thus, audition committees may actively advance women to the final round only when they are reasonably confident that the female candidate is above some threshold level of quality. If juries actively seek to increase the presence of women in the final round, they can do so only when there is no screen.

[42] As noted earlier, an obvious explanation for the importance of the individual fixed effects in the estimation is that the screen altered the pool of female applicants; however, we have been unable to show this empirically.

variates (in the even-numbered columns of Table 6). The inclusion of these individual covariates had little effect on the estimated effect of the screen.

A related concern is that those individuals who get hired at their first audition, and therefore do not contribute to the identification of the effect in the presence of individual fixed effects, are more able musicians than those who audition multiple times. (Alternatively, some individuals who audition and are not hired may get discouraged and not audition again and are therefore worse than those who audition multiple times.) Although this is a potential source of bias, it is important to remember that only a very small number of musicians win an audition in any given year, since there are just a handful of auditions (for a given instrument) among the major orchestras. Furthermore, many of the contestants in our sample did audition at least twice.

In addition, there are three pieces of empirical evidence that suggest this potential source of bias is not a major problem in our data. First, we control for the number of previous auditions in the even columns of Table 6, and this control does not change the results significantly. Second, there is no significant difference in the proportion female among those who auditioned both with and without a screen and those who auditioned only once (or who auditioned under only one policy regime). Finally, the coefficient estimates generated when the sample is restricted to those who auditioned at least three times are not perceptibly different from those generated from the full sample or from the sample of individuals who auditioned both with and without a screen. (These results are presented in Table A4.)

A third potential bias is that, because the effect of the screen is partially identified from differing audition procedures across orchestras, the results in Table 6 may indicate that orchestras that use screens are less discriminatory against women than those that do not. Specifically, because we include individual fixed effects, a bias would arise if women who are improving faster than average are more likely to audition for orchestras that use screens and are more likely to be advanced because these orchestras are intrinsically less discriminatory. Our sample contains only one orchestra per audition round that changed policy. As a result, we cannot separate the estimation by audition round and include orchestra fixed effects. We can, however, pool the audition rounds for the three orchestras that changed audition policy during our sample frame and include both individual and orchestra fixed effects.[43] These results are presented in Table 7.

In column (1) of Table 7 we include individual fixed effects, in which case the identification is from individuals who auditioned both with and without a screen. We add orchestra fixed effects in column (2) such that the identification now is from individuals who auditioned for a particular orchestra both before and after the orchestra began using a screen.[44] Finally, in column (3) we exclude individual but keep orchestra fixed effects to illustrate the importance of individual fixed effects. Again, the coefficient on "Blind" shows whether all musicians are more likely to be advanced when the audition is blind. The interaction between whether the individual is female and whether the audition is blind indicates whether women receive an extra boost relative to men when the screen is used.

The coefficient of interest is positive in columns (1) and (2) but negative in column (3), similar to the difference between the tabulations in Tables 4 and 5. In addition, the estimated effect of the blind auditions on the success of women is similar to that in Table 6. The point is that individual fixed-effects estimation matters; orchestra fixed effects, however, do not matter. In all cases, blind auditions increase the probability of advancement for both men and women. More

[43] We do not include the type of audition round since we have only one orchestra that changed procedures for the preliminaries, one that changed for the semifinals, and one that changed for the finals (and for which there were musicians who auditioned for that orchestra and audition round with and without a screen). We have also estimated these regressions separately for each of these three orchestras. Although the point estimates are not statistically significant, the magnitudes are quite similar to those presented in Table 6 for the corresponding round of the audition.

[44] In this subsample, there are 1,776 person-rounds comprised of individuals who auditioned for a particular orchestra, both behind and without a screen; on average these 552 individuals competed in 3.2 audition rounds.

TABLE 7—LINEAR PROBABILITY ESTIMATES OF THE LIKELIHOOD OF BEING ADVANCED: WITH INDIVIDUAL AND ORCHESTRA FIXED EFFECTS

	Include individual fixed effects		Exclude individual fixed effects
	(1)	(2)	(3)
Blind	0.404	0.399	0.103
	(0.027)	(0.027)	(0.018)
Female × Blind	0.044	0.041	−0.069
	(0.039)	(0.039)	(0.022)
Female			−0.005
			(0.019)
p-value of H_0: Blind + (Female × Blind) = 0	0.000	0.000	0.090
Individual fixed effects?	Yes	Yes	No
Orchestra fixed effects?	No	Yes	Yes
Year fixed effects?	Yes	Yes	Yes
Other covariates?	Yes	Yes	Yes
R^2	0.615	0.615	0.048
Number of observations	8,159	8,159	8,159

Notes: The unit of observation is a person-round. The dependent variable is 1 if the person is advanced to the next round and 0 if not. Standard errors are in parentheses. All specifications include an interaction for the sex being missing and a blind audition; "Other covariates" include automatic placement, years since last audition, number of auditions attended, size of the audition round, proportion female in audition round, whether a principal or substitute position, and a dummy indicating whether years since last audition is missing. These regressions include only the orchestras that changed their audition policy during our sample years and for which we observe individuals auditioning for the audition round both before and after the policy change. These regressions include 4,836 separate persons and are identified off of 1,776 person-rounds comprised of individuals who auditioned both before and after the policy change for a particular orchestra. *Source:* Eight-orchestra audition sample (three orchestras of which are used; see Notes). See text.

important, even though the effect is not statistically significant, the blind procedure has a positive effect on women's advancement.[45]

Finally, sex misclassification may also bias our estimates because, if the misclassification errors are uncorrelated with the equation error, the estimated effect of the screen will be attenuated (see, e.g., Richard Freeman, 1984). To address this potential problem, we use a less-subjective assessment of the probability that the individual is male or female. A U.S. Bureau of the Census tabulation, based on the postenumeration survey of the 1990 census, gives us the proportion female and male of the top 90 percent of all names.[46]

In Table 8 we estimate the same specifications given by columns (2), (4), (6), and (8) of Table 6 and column (2) of Table 7 using the census data in two ways. First, we simply replace our female covariate with the census probability.[47] Note that we also use a census estimate of the percentage of the audition round that is female (slightly changing our sample size), and a census estimate of the percentage of our sample for which the sex is indeterminate. In addition, our interaction term is constructed using the census probabilities. Second, we use

[45] Although the results from these three orchestras may not generalize to the other five, it should be noted that the coefficient estimate in column (3) of Table 7 is similar to that derived from a similar regression on the entire sample. This result is not surprising because the primary reason we are able to include both individual and orchestra fixed effects for these three orchestras is because they have unusually good record keeping, which allows us to observe the results of many auditions rather than another reason that might be correlated with how meritocratic the orchestra is.

[46] These data can be downloaded from http://www.census.gov/ftp/pub/genealogy/names. A possible problem with the data is that names are generational; a male name in one generation may become female in another.

[47] We do not impute census probabilities for the individuals whose sex we know with certainty (see footnote 31).

TABLE 8—LINEAR PROBABILITY ESTIMATES OF THE LIKELIHOOD OF BEING ADVANCED: ADDRESSING SEX MISCLASSIFICATION

Part A: Preliminary rounds

	Preliminaries			
	Without semifinals		With semifinals	
	OLS	IV	OLS	IV
Blind	−0.012	0.057	−0.174	0.290
	(0.043)	(0.045)	(0.093)	(0.241)
Female × Blind	0.139	0.137	0.272	−0.035
	(0.066)	(0.068)	(0.188)	(0.251)
Other covariates?	Yes	Yes	Yes	Yes
Individual fixed effects?	Yes	Yes	Yes	Yes
Year fixed effects?	Yes	Yes	Yes	Yes
R^2	0.771			
Number of observations	5,696	5,395	6,546	6,239

Part B: Semifinal and final rounds, and with orchestra fixed effects

	Semifinals		Finals		With orchestras fixed effects	
	OLS	IV	OLS	IV	OLS	IV
Blind	0.100	−0.197	−0.028	−0.025	0.010	0.061
	(0.083)	(0.700)	(0.125)	(0.141)	(0.028)	(0.033)
Female × Blind	−0.242	−0.193	0.160	0.324	0.069	0.052
	(0.120)	(0.429)	(0.171)	(0.181)	(0.035)	(0.036)
Other covariates	Yes	Yes	Yes	Yes	Yes	Yes
Individual fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.776		0.848		0.654	
Number of observations	1,600	1,360	1,509	1,127	8,882	8,159

Notes: The unit of observation is a person-round. The dependent variable is 1 if the individual is advanced to the next round and 0 if not. Standard errors are in parentheses. The instruments are the census probability that the individual is female, a dummy for whether the person has been sexed with certainty, and proportion female calculated using the census data and an interaction between whether the census data are missing and a screen has been used. The "OLS" columns use these as regressors. All specifications include an interaction for the sex being missing and a blind audition; "Other covariates" include automatic placement, years since last audition, number of auditions attended, whether a "Big Five" orchestra, size of the audition round, proportion female at the audition round, whether a principal or substitute position, and a dummy indicating whether years since last audition and automatic audition are missing. These are the same specifications as in columns (2), (4), (6), and (8) of Table 6 and column (2) of Table 7. The sample sizes change because in the even-numbered columns we simply replace our female covariate with the census probability and also use a census estimate of the percentage of the audition round that is female, which changes the sample size slightly.
Source: Eight-orchestra audition sample. See text.

the census probability as an instrument for our estimate (and for the percentage of the audition that is female, the percentage missing sex, and the interaction between female and whether the audition is blind).

The results are quite robust across these different methods for addressing potential measurement error. More important, the coefficients and their standard errors are generally similar in magnitude to those in Tables 6 and 7. With the exception of the semifinal round, the screen appears to have increased the likelihood that a woman would be advanced.[48]

[48] Another potential bias is from the short panel, which may affect the consistency of the estimates (Hsiao, 1986). We address the extent of this short panel problem in two ways. We first restrict our sample to those whom we observe auditioning at least three times (for the same round). Second, we restrict the estimation to those who auditioned at least once in a blind round and at least once in a not-blind round (those off of whom we are identified). The results do not change markedly from those in Table 6, showing that the short panel may not be a problem. See Table A4.

TABLE 9—LINEAR PROBABILITY ESTIMATES OF THE EFFECT OF BLIND AUDITIONS ON THE LIKELIHOOD OF BEING HIRED WITH INDIVIDUAL FIXED EFFECTS

	Without semifinals		With semifinals		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Completely blind audition	−0.024	0.047	0.001	0.006	0.001	0.005
	(0.028)	(0.041)	(0.009)	(0.011)	(0.008)	(0.009)
Completely blind audition × female	0.051	0.036	0.001	−0.004	0.011	0.006
	(0.046)	(0.048)	(0.016)	(0.016)	(0.013)	(0.013)
Year effects?	No	Yes	No	Yes	No	Yes
Other covariates?	No	Yes	No	Yes	No	Yes
R^2	0.855	0.868	0.692	0.707	0.678	0.691
Number of observations	4,108	4,108	5,883	5,883	9,991	9,991

Notes: The unit of observation is a person-round. The dependent variable is 1 if the individual is advanced (or hired) from the final round and 0 if not. Standard errors are in parentheses. All specifications include individual fixed effects, whether the sex is missing, and an interaction for sex being missing and a completely blind audition. "Other covariates" are the size of the audition, the proportion female at the audition, the number of individuals advanced (hired), whether a "Big Five" orchestra, the number of previous auditions, and whether the individual had an automatic semifinal or final.
Source: Eight-orchestra audition sample. See text.

C. *The Effect of the Screen on the Hiring of Women*

Using the Audition Sample.—Our analysis, thus far, has concerned the rounds of the audition process and the degree to which the screen enhances the likelihood of a woman's advancing from one round to the next. We turn now to the effect of the screen on the actual hire and estimate the likelihood an individual is hired out of the initial audition pool.[49] Whereas the use of the screen for each audition round was, more or less, an unambiguous concept, that for the entire process is not and we must define a blind audition. The definition we have chosen is that a blind audition contains all rounds that use the screen. In using this definition, we compare auditions that are completely blind with those that do not use the screen at all or use it for the early rounds only. We divide the sample into auditions that have a semifinal round and those that do not, because the previous analysis suggested they might differ.

The impact of completely blind auditions on the likelihood of a woman's being hired is given in Table 9, for which all results include individual fixed effects.[50] The impact of the screen is positive and large in magnitude, but only when there is no semifinal round. Women are about 5 percentage points more likely to be hired than are men in a completely blind audition, although the effect is not statistically significant. The effect is nil, however, when there is a semifinal round, perhaps as a result of the unusual effects of the semifinal round. The impact for all rounds [columns (5) and (6)] is about 1 percentage point, although the standard errors are large and thus the effect is not statistically significant. Given that the probability of winning an audition is less than 3 percent, we would need more data than we currently have to estimate a statistically significant effect, and even a 1-percentage-point increase is large, as we later demonstrate.

[49] There are four auditions in which the committee could not choose between two players and therefore asked each to play with the orchestra. We consider both to be winners. The results are not sensitive to this classification. For this analysis we exclude auditions with no women, all women, or no winner; these exclusions do not change the results.

[50] In Table 9 we are identified off of individuals who competed in auditions that were completely blind *and* those that were not completely blind (that is, *any* one round could not be blind). The unit of observation is the person-round and there are 92 fulfilling this criterion for auditions without a semifinal [columns (1) and (2)]; on average these persons competed in 3.6 auditions in this sample. There are 625 person-rounds fulfilling this criterion that included a semifinal [columns (3) and (4)] and on average these persons competed in 3.5 auditions in this sample. Finally, there are 911 person-rounds fulfilling this criterion across all audition [columns (5) and (6)] and on average these persons competed in 3.5 auditions in this sample. The sample off of which we are identified is larger for all auditions than for the sum of the other two because some individuals auditioned both with and without a semifinal round.

TABLE 10—PROBIT ESTIMATES OF THE EFFECT OF BLIND AUDITIONS ON THE SEX OF NEW MEMBERS: 1970 TO 1996

	Any blind auditions	Only blind preliminaries and/or semifinals vs. completely blind auditions
	(1)	(2)
Any blind auditions	0.238	
	(0.183)	
	[0.075]	
Only blind preliminaries and/or semifinals		0.232
		(0.184)
		[0.074]
Completely blind auditions		0.361
		(0.438)
		[0.127]
Section:		
Woodwinds	−0.187	−0.188
	(0.114)	(0.114)
	[−0.058]	[−0.058]
Brass	−1.239	−1.237
	(0.157)	(0.157)
	[−0.284]	[−0.284]
Percussion	−1.162	−1.164
	(0.305)	(0.305)
	[−0.235]	[−0.235]
p-value of test: only blind preliminaries and/or semifinals = completely blind		0.756
pseudo R^2	0.106	0.106
Number of observations	1,128	1,128

Notes: The dependent variable is 1 if the individual is female and 0 if male. Standard errors are in parentheses. All specifications include orchestra fixed effects and orchestra-specific time trends. Changes in probabilities are in brackets; see text for an explanation of how they are calculated. New members are those who enter the orchestra for the first time. Returning members are not considered new. The omitted section is strings.
Source: Eleven-orchestra roster sample. See text.

Using the Roster Data.—The roster data afford us another way to evaluate the effect of the screen on the sex composition of orchestras. Using the rosters we know the sex of new hires each year for 11 orchestras, and we also have information (see Table 1) on the year the screen was adopted by each orchestra. We treat the orchestra position as the unit of observation and ask whether the screen affects the sex of the individual who fills the position. We model the likelihood that a female is hired in a particular year as a function of whether the orchestra's audition procedure involved a screen, again relying on the variation over time within a particular orchestra. Thus, in all specifications, we include orchestra fixed effects and an orchestra-specific time trend.

The roster data extend further back in time than do the audition data and could conceivably begin with the orchestra's founding, although there is no obvious reason to include many years when none used the screen. We report, in Table 10, the effects of the screen on the hiring of women from 1970 to 1996 using a probit model. The screen is first defined to include any blind auditions [column (1)]. In column (2) we estimate separate effects for orchestras using blind preliminary (and semifinal) rounds but not blind finals and those with completely blind auditions.

To interpret the probit coefficient, we first predict a base probability, under the assumption that each orchestra does not use a screen. We then predict a new probability assuming the orchestra uses a screen. The mean difference in the probabilities is given in brackets.

The coefficient on blind in column (1) is positive, although not significant at any usual level of confidence. The estimates in column (2) are positive and equally large in magnitude to those in column (1). Further, these estimates show that the existence of any blind round makes a difference and that a completely blind process has a somewhat larger effect (albeit with a large standard error).[51] According to the point estimates in column (1) of Table 10, blind auditions increase the likelihood a female will be hired by 7.5 percentage points. The magnitude of the effect must be judged relative to the overall average and, for the period under consideration, it was about 30 percent.[52] Thus blind auditions increased the likelihood a female would be hired by 25 percent.

Making Further Sense of the Results on Hiring.—The audition sample results suggest that blind auditions increase the probability of eventual success for a female candidate by 5 percentage points, but only if there is no semifinal round. The average effect for both types of auditions is closer to 1 percentage point (with a large standard error). The following example, using assumed values based on the actual data, demonstrates that an increase of about 2 percentage points in the probability of a woman's success out of an audition can explain the entire change in female hires, allowing the share of candidates who are female to increase from 0.2 to 0.3. Thus an increase of 1 percentage point—our point estimate—can account for a substantial share.

Consider two regimes: one without the screen (not blind) and another with the screen (blind). In the not-blind regime, assume that 20 percent of the candidates are female and that in the blind regime 30 percent are female.[53] We know that in the era (say, before 1970) when few orchestras used the screen for the preliminary round (see Table 1), 10 percent (that is, 0.0996) of new hires were women. Also assume that 30 candidates enter each audition, independent of audition regime, and that one musician is hired out of each audition. Using these assumptions, taken from the actual data, the success rate for the typical female audition candidate in the not-blind regime will be 0.0166 and that for the typical male will be 0.0375. If in the blind regime, however, the percentage of new hires who are female increases to 35 percent (its approximate figure for the past 10 years), the success rate for a female audition candidate must have increased to 0.0389 (and that for a male must have decreased to 0.0310). That is, for consistency with the data on percent female, the success rate for female candidates would have had to increase by about 2.2 percentage points, moving from the not-blind to the blind regime. Our point estimate is that about half of that increase—1 percentage point—was the result of the effect of the screened audition process.

Using the example we just offered, the increase in the probability of a woman's being hired out of an audition accounts for 66 percent of the total increase in the fraction female among new hires. Half of the 66 percent comes from the switch to blind auditions.[54] The other half could have resulted, for example, from a

[51] We have also attempted to interact the effect of blind auditions with section dummies. We find that the main effect of blind auditions is almost identical to that for the string section, which is not surprising given that the strings comprise 65 percent of the observations. In addition, fewer than 4 percent of the musicians hired into the percussion and brass sections are female.

[52] See Table A2.

[53] The fraction female in the not-blind regime (taking it to be the period before 1970) is 0.187 in our data (see Table 3). In the blind regime it was between 0.35 and 0.4. We have chosen the more conservative 0.3 in the example because we want to use a number that is independent of the switch to using the screen. That is, we would like to use a fraction female that is solely the result of increases in female participation in general but independent of changes in audition procedures.

[54] The proportion female among new hires is $(n \cdot \lambda \cdot \alpha)$, where n = the number of audition candidates (in this example $n = 30$); λ = the success rate of the average female candidate, which may be enhanced by the screen (in this example λ increases from 0.0166 to 0.0389 or by 2.2 percentage points, about half of which is due to the screen, based on our estimates); and α = the fraction female among candidates (assumed here to increase from 20 to 30 percent independent of λ). The percentage of the total change accounted for by the change in λ is given by $(n \cdot \alpha \cdot \Delta\lambda)/\Delta(n \cdot \lambda \cdot \alpha)$ or on average by $[(30 \cdot 0.25 \cdot 0.022)/(0.35 - 0.0996) = 66$ percent. (The 0.25 figure is the average of that in the treatment period and that previously.) Since half is accounted for by the screen, about 33 percent of the increase in the proportion female among new hires comes from the blind audition process.

greater acceptance of female musicians by music directors. The remainder (34 percent) of the increase in the fraction female among new hires is accounted for by the increased percentage female among audition candidates. That portion comes primarily from the increase in the fraction female among music school graduates.

The point estimates from the roster data also suggest that a substantial portion of the increase in female hires across the two regimes, not-blind and blind, can be explained by the change in audition procedures. In the not-blind regime about 10 percent of all hires are female but in the blind regime about 35 percent are, a difference of 25 percentage points. The estimates in column (1) of Table 10 show that the switch to the blind regime increases the likelihood a woman will be hired by 7.5 percentage points—30 percent of the total change—although we emphasize that the coefficient is imprecisely estimated.

One may wonder why there was disparate treatment of female musicians before the screen was used. A great orchestra is not simply a collection of the finest musicians. It is, rather, a group of great musicians who play magnificently as an ensemble. Substantial amounts of specific human capital are acquired on the job and tenure differences by sex, therefore, could influence hiring decisions.[55] Leaves of absence are ordinarily allowed for medical (including maternity) and professional reasons. We find, using the roster sample from 1960 to 1996, that the average female musician took 0.067 leaves per year, whereas the average male musician took 0.061, a difference that is not statistically significant, and that their length of leave was trivially different. Tenure differences were also small and some specifications show that women accumulated more years with an orchestra, given their starting year and orchestra.[56] Turnover and leaves of absence do not appear to differ by sex and thus should not have rationally influenced hiring decisions.

IV. Conclusion

The audition procedures of the great U.S. symphony orchestras began to change sometime in the 1970's. The changes included increasing the number of candidates at auditions—a democratization of the process—and using a physical screen during the audition to conceal the candidate's identity and ensure impartiality. We analyze what difference blind auditions have meant for female musicians.

We have collected, from orchestral management files and archives, a sample of auditions for eight major orchestras. These records contain the names of all candidates and identify those advanced to the next round, including the ultimate winner of the competition. The data provide a unique means of testing whether discrimination existed in the various rounds of a hiring process and even allow the linkage of individuals across auditions. A strong presumption exists that discrimination has limited the employment of female musicians, especially by the great symphony orchestras. Not only were their numbers extremely low until the 1970's, but many music directors, ultimately in charge of hiring new musicians, publicly disclosed their belief that female players had lower musical talent.

The question is whether hard evidence can support an impact of discrimination on hiring. Our analysis of the audition and roster data indicates that it can, although we mention various caveats before we summarize the reasons. Even though our sample size is large, we identify the coefficients of interest from a much smaller sample. Some of our coefficients of interest, therefore, do not pass standard tests of statistical significance and there is, in addition, one persistent result that goes in the opposite direction. The weight of the evidence, however, is what we find most persuasive and what we

[55] Musicians of the Vienna Philharmonic made this argument in a radio broadcast by the West German State Radio in February 1996 [translation provided by William Osborne]. See also *New York Times* (1996) in which a player for the Vienna Philharmonic argued that female musicians would cost the orchestra considerably more because substitutes would have to be hired if they became pregnant.

[56] The general specification is number of actual years with an orchestra as a function of the starting year, section dummies, and a female dummy, for the period since 1959. The coefficient on the female dummy is −0.299 with a large standard error (the mean of tenure is 11.7 years). With the addition of orchestra fixed effects, the coefficient on the female dummy is +0.062, again with a large standard error. The difference in tenure by sex, therefore, is extremely small.

have emphasized. The point estimates, moreover, are almost all economically significant.

Using the audition data, we find that the screen increases—by 50 percent—the probability that a woman will be advanced from certain preliminary rounds and increases by severalfold the likelihood that a woman will be selected in the final round. By the use of the roster data, the switch to blind auditions can explain 30 percent of the increase in the proportion female among new hires and possibly 25 percent of the increase in the percentage female in the orchestras from 1970 to 1996.[57] As in research in economics and other fields on double-blind refereeing (see, e.g., Blank, 1991), the impact of a blind procedure is toward impartiality and the costs to the journal (here to the orchestra) are relatively small. We conclude that the adoption of the screen and blind auditions served to help female musicians in their quest for orchestral positions.

[57] The point estimate for the increased likelihood a woman would be a new hire, as a result of the adoption of blind auditions, is 7.5 percentage points using the roster data (see Table 10). Because the percentage female among new hires increased from 10 to 35 percent from before 1970 to the 1990's, our estimate implies that 30 percent of the 25 percentage-point increase can be explained by the adoption of the screen. How this increase affected the percentage female in the orchestra depends on the sex composition of the orchestra, retirement (or turnover), and the time frame. We assume a 25-year time frame (from 1970 to 1995) and two retirements (thus two hires) per year. An increase in the percentage female among new hires from 10 percent (its level pre-1970) to 17.5 percent (10 + 7.5%) implies that in 25 years, 13.75 women (out of 100) will be in the orchestra, or an increase of 3.75. The actual increase was 15 women, meaning 25 percent of the increase can be explained by the adoption of the screen. We assume in this example that the age distribution of the 100 players in 1970 is uniform between ages 25 and 74, that all hires occur at age 25, and that men and women are drawn from the same age distribution.

Appendix

Table A1—Sample Descriptive Statistics, Audition Data

	Preliminaries							
	Without semifinals		With semifinals		Semifinals		Finals	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Advanced	0.184	0.387	0.185	0.388	0.349	0.477	0.200	0.400
Blind	0.793	0.405	0.976	0.152	0.808	0.394	0.122	0.328
Female	0.376	0.485	0.374	0.484	0.410	0.492	0.411	0.492
Female × Blind	0.305	0.461	0.362	0.481	0.325	0.469	0.056	0.230
Missing female	0.002	0.047	0.002	0.047	0.004	0.066	0	0
Missing female × Blind	0.002	0.043	0.002	0.047	0.004	0.061	0	0
Years since last audition	2.480	1.661	2.621	2.209	2.432	2.393	2.272	1.895
Years since last audition, missing	0.663	0.473	0.505	0.500	0.386	0.487	0.505	0.500
Automatic placement	—	—	—	—	0.267	0.443	0.137	0.345
Number of auditions attended	1.611	1.137	2.147	1.717	2.490	1.886	2.051	1.513
"Big Five" orchestra	0.607	0.488	0.323	0.467	0.213	0.409	0.391	0.488
Total number of auditioners	44.348	22.202	64.279	35.914	15.054	7.187	8.622	4.445
Proportion female at round	0.375	0.206	0.373	0.239	0.407	0.211	0.411	0.213
Principal	0.192	0.394	0.368	0.482	0.353	0.478	0.278	0.448
Substitute	0.025	0.157	0.005	0.071	0.010	0.101	0.021	0.141
Number of observations (person-rounds)	5,395		6,239		1,360		1,127	

Source: Eight-orchestra audition sample. See text.

TABLE A2—SAMPLE DESCRIPTIVE STATISTICS, ROSTER DATA: 1970 TO 1996

	Mean	Standard deviation
Proportion female among new hires	0.293	0.455
$(\text{Proportion female})_{t-1}$	0.179	0.081
Only blind preliminary auditions	0.572	0.495
All auditions blind	0.104	0.305
Section:		
Strings	0.642	0.480
Woodwinds	0.158	0.365
Brass	0.165	0.371
Percussion	0.035	0.185
Number of observations	1,128	

Note: Means are musician weighted, not audition weighted.
Source: Eleven-orchestra roster sample. See text.

TABLE A3—LINEAR PROBABILITY ESTIMATES OF THE LIKELIHOOD OF BEING ADVANCED: BY ROUND

	Preliminaries							
	Without semifinals		With semifinals		Semifinals		Finals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.007	0.011	−0.054	−0.085	0.103	0.099	0.002	0.0004
	(0.025)	(0.025)	(0.069)	(0.069)	(0.061)	(0.061)	(0.028)	(0.028)
Female × Blind	−0.062	−0.067	0.005	0.037	−0.142	−0.137	−0.091	−0.078
	(0.028)	(0.028)	(0.070)	(0.070)	(0.066)	(0.067)	(0.075)	(0.075)
Blind audition	0.015	0.040	0.024	0.027	0.053	0.115	0.058	0.123
	(0.022)	(0.030)	(0.057)	(0.062)	(0.049)	(0.078)	(0.058)	(0.089)
p-value of H_0: Female + (Female × Blind) = 0	0.000	0.000	0.000	0.000	0.210	0.222	0.207	0.271
Other covariates?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrument fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Orchestra fixed effects?	No	Yes	No	Yes	No	Yes	No	Yes
R^2	0.062	0.070	0.033	0.045	0.074	0.081	0.064	0.068
Number of observations (person-rounds)	5,395	5,395	6,239	6,239	1,360	1,360	1,127	1,127

Notes: The dependent variable is 1 if the individual is advanced to the next round and 0 if not. Standard errors are in parentheses. All specifications include dummies indicating whether the sex is missing, and an interaction for the sex being missing and a blind audition. "Other covariates" include automatic round, number of auditions attended, whether a "Big Five" orchestra, size of round, proportion female at the round, and whether a principal (including assistant and associate principal) or substitute position; except in columns (2), (4), (6), and (8) for which "Other covariates" include only automatic placement and number of auditions attended. These results are comparable to those in Table 6 but without individual fixed effects.
Source: Eight-orchestra audition sample. See text.

TABLE A4—LINEAR PROBABILITY ESTIMATES OF THE LIKELIHOOD OF BEING ADVANCED: ASSESSING SHORT-PANEL BIAS

	Part A: Preliminary rounds			
	Preliminaries			
	Without semifinals		With semifinals	
	I[a]	II[b]	I[a]	II[b]
Blind	−0.024	−0.042	−0.047	−0.095
	(0.066)	(0.063)	(0.383)	(0.744)
Female × Blind	0.126	0.095	−0.035	0.041
	(0.095)	(0.071)	(0.403)	(0.275)
p-value of H_0: Blind + (Female × Blind) = 0	0.233	0.502	0.807	0.943
Other covariates?	Yes	Yes	Yes	Yes
Individual fixed effects?	Yes	Yes	Yes	Yes
Year fixed effects?	Yes	Yes	Yes	Yes
R^2	0.491	0.537	0.423	0.732
Number of observations (person-rounds)	1,025	639	1,928	55

	Part B: Semifinals and finals, and with orchestra fixed effects					
	Semifinals		Finals		With orchestras fixed effects	
	I[a]	II[b]	I[a]	II[b]	I[a]	II[b]
Blind	0.060	0.169	0.123	−0.140	0.084	0.352
	(0.133)	(0.109)	(0.356)	(0.449)	(0.047)	(0.056)
Female × Blind	−0.179	−0.284	0.157	0.403	0.042	0.021
	(0.195)	(0.142)	(0.408)	(0.415)	(0.051)	(0.041)
p-value of H_0: Blind + (Female × Blind) = 0	0.438	0.298	0.212	0.587	0.011	0.000
Other covariates?	Yes	Yes	Yes	Yes	Yes	Yes
Individual fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.438	0.593	0.721	0.728	0.506	0.603
Number of observations (person-rounds)	269	223	127	67	2,321	1,776

Notes: The dependent variable is 1 if the individual is advanced to the next round and 0 if not. Standard errors are in parentheses. These are the same specifications as in columns (2), (4), (6), and (8) of Table 6 and column (2) of Table 7.
Source: Eight-orchestra audition sample. See text.

[a] Includes those who auditioned at least three times (for the relevant round).

[b] Includes those who auditioned at least once in a blind audition and at least once in a not-blind audition (for the relevant round).

REFERENCES

Allmendinger, Jutta; Hackman, J. Richard and Lehman, Erin V. "Life and Work in Symphony Orchestras." *Music Quarterly*, Summer 1996, *80*(2), pp. 194–219.

Arrow, Kenneth. "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in labor markets*. Princeton, NJ: Princeton University Press, 1973, pp. 3–33.

Ayres, Ian and Waldfogel, Joel. "A Market Test for Race Discrimination in Bail Setting." *Stanford Law Review*, May 1994, *46*(5), pp. 987–1047.

Becker, Gary. *The economics of discrimination*, 2nd Ed. Chicago: University of Chicago Press, 1971 [orig. pub. 1957].

Blank, Rebecca. "The Effects of Double-Blind versus Single-Blind Refereeing: Experimental Evidence from the *American Economic Review*." *American Economic Review*, December 1991, *81*(5), pp. 1041–67.

Economist, The. "American Orchestras: All Ears." November 30, 1996.

Ellefson, Connie Lockhard. *The melting pot book of baby names*, 2nd Ed. Cincinnati, OH: Better Way Books, 1990.

England, Paula. "The Failure of Human Capital Theory to Explain Occupational Sex Segregation." *Journal of Human Resources*, Summer 1982, *17*(3), pp. 358–70.

England, Paula; Kilbourne, Barbara Stanek; Farkas, George and Dou, Thomas. "Explaining Occupational Sex Segregation and Wages: Findings from a Model with Fixed Effects." *American Sociological Review*, August 1988, *53*(4), pp. 544–58.

Filer, Randall K. "Occupational Segregation, Compensating Differentials, and Comparable Worth," in Robert T. Michael, Heidi I. Hartmann, and Brigit O'Farrell, eds., *Pay equity: Empirical inquiries*. Washington, DC: National Academy Press, 1989, pp. 153–70.

Freeman, Richard. "Longitudinal Analyses of the Effects of Trade Unions." *Journal of Labor Economics*, January 1984, *2*(1), pp. 1–26.

Holzer, Harry and Neumark, David. "Are Affirmative Action Hires Less Qualified? Evidence from Employer-Employee Data on New Hires." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5603, June 1996.

Hsiao, Cheng. *Analysis of panel data*. New York: Cambridge University Press, 1986.

Kenney, Genevieve and Wissoker, Douglas A. "An Analysis of the Correlates of Discrimination Facing Young Hispanic Job-Seekers." *American Economic Review*, June 1994, *84*(3), pp. 674–83.

Kolatch, Alfred J. *The Jonathan David dictionary of first names*. New York: Perigee Books, 1990.

Lansky, Bruce. *35,000+ baby names*. New York: Meadowbrook Press, 1995.

Neumark, David (with Bank, Roy and Van Nort, Kyle D.). "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics*, August 1996, *111*(3), pp. 915–41.

New York Times. "Vienna Boys' Choir Hands the Baton to a Woman." October 16, 1996.

Polachek, Solomon W. "Occupational Segregation among Women: Theory, Evidence and a Prognosis," in Cynthia B. Lloyd, Emily Andrews, and Curtis Gilroy, eds., *Women in the labor market*. New York: Columbia University Press, 1979, pp. 137–57.

Seltzer, George. *Music matters: The performer and the American Federation of Musicians*. Metuchen, NJ: Scarecrow Press, 1989.

This article has been cited by:

1. Zara Whysall. Cognitive Biases in Recruitment, Selection, and Promotion: The Risk of Subconscious Discrimination 215-243. [Crossref]
2. Kristin J. Kleinjans, Karl Fritjof Krassel, Anthony Dukes. 2017. Occupational Prestige and the Gender Wage Gap. *Kyklos* **70**:4, 565-593. [Crossref]
3. Heather Savigny. 2017. Cultural Sexism is Ordinary: Writing and Re-Writing Women in Academia. *Gender, Work & Organization* **24**:6, 643-655. [Crossref]
4. Michael R. Strain, Douglas A. Webber. 2017. High school experiences, the gender wage gap, and the selection of occupation. *Applied Economics* **49**:49, 5040-5049. [Crossref]
5. Sam de Boise. 2017. Gender Inequalities and Higher Music Education: Comparing the UK and Sweden. *British Journal of Music Education* **20**, 1-19. [Crossref]
6. Georgios Georgiou. 2017. Are oral examinations objective? Evidence from the hiring process for judges in Greece. *European Journal of Law and Economics* **44**:2, 217-239. [Crossref]
7. Guido Cozzi, Silvia Galli, Noemi Mantovan. 2017. Will a Shrink Make you Richer? Gender Differences in the Effects of Psychotherapy on Labour Efficiency. *European Economic Review* . [Crossref]
8. Rocco Palumbo, Reginald B. Adams, Ursula Hess, Robert E. Kleck, Leslie Zebrowitz. 2017. Age and Gender Differences in Facial Attractiveness, but Not Emotion Resemblance, Contribute to Age and Gender Stereotypes. *Frontiers in Psychology* **8**. . [Crossref]
9. Francine D. Blau, Lawrence M. Kahn. 2017. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature* **55**:3, 789-865. [Abstract] [View PDF article] [PDF with links]
10. Kristen Roche. 2017. Millennials and the Gender Wage Gap in the U.S.: A Cross-Cohort Comparison of Young Workers Born in the 1960s and the 1980s. *Atlantic Economic Journal* **45**:3, 333-350. [Crossref]
11. Lily Hua Fang, Sterling Huang. 2017. Gender and Connections among Wall Street Analysts. *The Review of Financial Studies* **30**:9, 3305-3335. [Crossref]
12. Danny Cohen-Zada, Alex Krumer, Mosi Rosenboim, Offer Moshe Shapir. 2017. Choking under pressure and gender: Evidence from professional tennis. *Journal of Economic Psychology* **61**, 176-190. [Crossref]
13. Juan Manuel Garc?a Lara, Beatriz Garc?a Osma, Araceli Mora, Mariano Scapin. 2017. The monitoring role of female directors over accounting quality. *Journal of Corporate Finance* **45**, 651-668. [Crossref]
14. David B. Bills, Valentina Di Stasio, Klarita Gërxhani. 2017. The Demand Side of Hiring: Employers in the Labor Market. *Annual Review of Sociology* **43**:1, 291-310. [Crossref]
15. Ulf Nielsson, Herdis Steingrimsdottir. 2017. The signalling value of education across genders. *Empirical Economics* **8**. . [Crossref]
16. Paulo Júlio, José Tavares. 2017. The Good, the Bad and the Different: Can Gender Quotas Raise the Quality of Politicians?. *Economica* **84**:335, 454-479. [Crossref]
17. Jason Chan, Jing Wang. 2017. Hiring Preferences in Online Labor Markets: Evidence of a Female Hiring Bias. *Management Science* . [Crossref]
18. Aaron Meskin, Jon Robson, Anna Ichino, Kris Goffin, Annelies Monser?. 2017. Philosophical aesthetics and cognitive science. *Wiley Interdisciplinary Reviews: Cognitive Science* **71**, e1445. [Crossref]
19. Petr Houdek. 2017. Is Behavioral Ethics Ready for Giving Business and Policy Advice?. *Journal of Management Inquiry* 105649261771289. [Crossref]

20. Shaun M. Dougherty, Joshua S. Goodman, Darryl V. Hill, Erica G. Litke, Lindsay C. Page. 2017. Objective course placement and college readiness: Evidence from targeted middle school math acceleration. *Economics of Education Review* **58**, 141-161. [Crossref]
21. Sean Hermanson. 2017. Implicit Bias, Stereotype Threat, and Political Correctness in Philosophy. *Philosophies* **2**:2, 12. [Crossref]
22. Jason Greenberg, Ethan Mollick. 2017. Activist Choice Homophily and the Crowdfunding of Female Founders. *Administrative Science Quarterly* **62**:2, 341-374. [Crossref]
23. Helen F. Mitchell. 2017. Music students' perceptions of experiential learning at the moot audition. *Music Education Research* **156**, 1-12. [Crossref]
24. Helen Mitchell, Roger Benedict. 2017. The moot audition: Preparing music performers as expert listeners. *Research Studies in Music Education* **2**, 1321103X1770963. [Crossref]
25. Seth Carnahan, Brad N. Greenwood. 2017. Managers' Political Beliefs and Gender Inequality among Subordinates: Does His Ideology Matter More Than Hers?. *Administrative Science Quarterly* **84**, 000183921770878. [Crossref]
26. Alberto Bacchelli, Moritz Beller. Double-Blind Review in Software Engineering Venues: The Community's Perspective 385-396. [Crossref]
27. Aaron Williamon, George Waddell. 2017. Eye of the Beholder: Stage Entrance Behavior and Facial Expression Affect Continuous Quality Ratings in Music Performance. *Frontiers in Psychology* **8**. . [Crossref]
28. Paul Gift. 2017. Performance Evaluation and Favoritism. *Journal of Sports Economics* **14**, 152700251770242. [Crossref]
29. Brian R. Hedden. 2017. Should juries deliberate?. *Social Epistemology* **87**, 1-19. [Crossref]
30. Elisabetta Addis, Majlinda Joxhe. 2017. Gender Gaps in Social Capital: A Theoretical Interpretation of Evidence from Italy. *Feminist Economics* **23**:2, 146-171. [Crossref]
31. Shane Thompson. 2017. COLLEGE ADVISING AND GENDER. *Economic Inquiry* **55**:2, 1007-1016. [Crossref]
32. Linda Babcock, Maria P. Recalde, Lise Vesterlund, Laurie Weingart. 2017. Gender Differences in Accepting and Receiving Requests for Tasks with Low Promotability. *American Economic Review* **107**:3, 714-747. [Abstract] [View PDF article] [PDF with links]
33. Joya Misra, Laurel Smith-Doerr, Nilanjana Dasgupta, Gabriela Weaver, Jennifer Normanly. 2017. Collaboration and Gender Equity among Academic Scientists. *Social Sciences* **6**:1, 25. [Crossref]
34. Simon Moss, S. Ram Vemuri, Darren Hedley, Mirko Uljarevic. 2017. Initiatives that diminish the biases of recruiters against people who disclose or demonstrate a diagnosed mental disorder. *Equality, Diversity and Inclusion: An International Journal* **36**:1, 2-16. [Crossref]
35. Robert J. MacCoun, Saul Perlmutter. Blind Analysis as a Correction for Confirmatory Bias in Physics and in Psychology 295-322. [Crossref]
36. Raymond Fisman, Daniel Paravisini, Vikrant Vig. 2017. Cultural Proximity and Loan Outcomes. *American Economic Review* **107**:2, 457-492. [Abstract] [View PDF article] [PDF with links]
37. Sandra Ludwig, Gerlinde Fellner-Röhling, Carmen Thoma. 2017. Do women have more shame than men? An experiment on self-assessment and the shame of overestimating oneself. *European Economic Review* **92**, 31-46. [Crossref]
38. Thomas de Haan, Theo Offerman, Randolph Sloof. 2017. Discrimination in the Labour Market: The Curse of Competition between Workers. *The Economic Journal* **100**. . [Crossref]

39. Jørgen Juel Andersen, Niels Johannesen, David Dreyer Lassen, Elena Paltseva. 2017. Petro Rents, Political Institutions, and Hidden Wealth: Evidence from Offshore Bank Accounts. *Journal of the European Economic Association* . [Crossref]
40. Andrea C. Tricco, Sonia M. Thomas, Jesmin Antony, Patricia Rios, Reid Robson, Reena Pattani, Marco Ghassemi, Shannon Sullivan, Inthuja Selvaratnam, Cara Tannenbaum, Sharon E. Straus. 2017. Strategies to Prevent or Reduce Gender Bias in Peer Review of Research Grants: A Rapid Scoping Review. *PLOS ONE* **12**:1, e0169718. [Crossref]
41. Iñaki R Longarela. 2017. Explaining vertical gender segregation: a research agenda. *Work, employment and society* **13**, 095001701666813. [Crossref]
42. Angela L. Workman-Stark. Barriers to Inclusion 61-82. [Crossref]
43. M. Bertrand, E. Duflo. Field Experiments on Discrimination a aLaura Stilwell and Jan Zilinsky provided excellent research assistance. We thank Abhijit Banerjee for comments. We are particularly grateful to Betsy Levy Paluck, our discussant, for her detailed and thoughtful review of an earlier draft 309-393. [Crossref]
44. Eva Derous, Jeroen Decoster. 2017. Implicit Age Cues in Resumes: Subtle Effects on Hiring Discrimination. *Frontiers in Psychology* **8**. . [Crossref]
45. CHARLES L. BALLARD, JEREMY F. DUFF. 2017. The Effects of Perceptions of Economic Inequality on Policy Preferences: Evidence from Michigan. *Michigan Academician* **44**:3, 256-286. [Crossref]
46. Darius Palia. 2016. Differential Access to Capital from Financial Institutions by Minority Entrepreneurs. *Journal of Empirical Legal Studies* **13**:4, 756-785. [Crossref]
47. Johnson B. Lightfoote, Curtiland Deville, Loralie D. Ma, Karen M. Winkfield, Katarzyna J. Macura. 2016. Diversity, Inclusion, and Representation: It Is Time to Act. *Journal of the American College of Radiology* **13**:12, 1421-1425. [Crossref]
48. Isabel Fernandez-Mateo, Roberto M. Fernandez. 2016. Bending the Pipeline? Executive Search and Gender Inequality in Hiring for Top Management Jobs. *Management Science* **62**:12, 3636-3655. [Crossref]
49. STEVEN ONGENA, ALEXANDER POPOV. 2016. Gender Bias and Credit Access. *Journal of Money, Credit and Banking* **48**:8, 1691-1724. [Crossref]
50. Daphne Brandenburg. 2016. Implicit attitudes and the social capacity for free will. *Philosophical Psychology* **29**:8, 1215-1228. [Crossref]
51. Amanda L. Golbeck, Arlene Ash, Mary Gray, Marcia Gumpertz, Nicholas P. Jewell, Jon R. Kettenring, Judith D. Singer, Yulia R. Gel. 2016. A conversation about implicit bias. *Statistical Journal of the IAOS* **32**:4, 739-755. [Crossref]
52. Martin L. Jönsson, Julia Sjödahl. 2016. INCREASING THE VERACITY OF IMPLICITLY BIASED RANKINGS. *Episteme* 1-19. [Crossref]
53. Krista Hyde. 2016. Testimonial Injustice and Mindreading. *Hypatia* **31**:4, 858-873. [Crossref]
54. Amanda Bayer, Cecilia Elena Rouse. 2016. Diversity in the Economics Profession: A New Attack on an Old Problem. *Journal of Economic Perspectives* **30**:4, 221-242. [Abstract] [View PDF article] [PDF with links]
55. Guillaume Beaurain, David Masclet. 2016. Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence. *European Economic Review* **90**, 350-362. [Crossref]
56. María Arrazola, José de Hevia. 2016. The Gender Wage Gap in Offered, Observed, and Reservation Wages for Spain. *Feminist Economics* **22**:4, 101-128. [Crossref]

57. Michèle Céline Kaufmann, Franciska Krings, Sabine Sczesny. 2016. Looking Too Old? How an Older Age Appearance Reduces Chances of Being Hired. *British Journal of Management* **27**:4, 727-739. [Crossref]
58. Fernanda Brollo, Ugo Troiano. 2016. What happens when a woman wins an election? Evidence from close races in Brazil. *Journal of Development Economics* **122**, 28-45. [Crossref]
59. Samuel R. Lucas. 2016. Where the Rubber Meets the Road. *Socius: Sociological Research for a Dynamic World* **2**, 237802311663470. [Crossref]
60. Alexander Buijsrogge, Eva Derous, Wouter Duyck. 2016. Often biased but rarely in doubt: How initial reactions to stigmatized applicants affect interviewer confidence. *Human Performance* **29**:4, 275-290. [Crossref]
61. Jan Feld, Nicolás Salamanca, Daniel S. Hamermesh. 2016. Endophilia or Exophobia: Beyond Discrimination. *The Economic Journal* **126**:594, 1503-1527. [Crossref]
62. Carol Bates, Lynn Gordon, Elizabeth Travis, Archana Chatterjee, Linda Chaudron, Barbara Fivush, Martha Gulati, Reshma Jagsi, Poonam Sharma, Marin Gillis, Rebecca Ganetzky, Amelia Grover, Diana Lautenberger, Ashleigh Moses. 2016. Striving for Gender Equity in Academic Medicine Careers. *Academic Medicine* **91**:8, 1050-1052. [Crossref]
63. Oriel FeldmanHall, Tim Dalgleish, Davy Evans, Lauren Navrady, Ellen Tedeschi, Dean Mobbs. 2016. Moral Chivalry. *Social Psychological and Personality Science* **7**:6, 542-551. [Crossref]
64. Thomas Breda, Mélina Hillion. 2016. Teaching accreditation exams reveal grading biases favor women in male-dominated disciplines in France. *Science* **353**:6298, 474-478. [Crossref]
65. Vojtěch Bartoš, Michal Bauer, Julie Chytilová, Filip Matějka. 2016. Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition. *American Economic Review* **106**:6, 1437-1475. [Abstract] [View PDF article] [PDF with links]
66. Alessandro Acquisti, Curtis Taylor, Liad Wagman. 2016. The Economics of Privacy. *Journal of Economic Literature* **54**:2, 442-492. [Abstract] [View PDF article] [PDF with links]
67. Diana L. Miller. 2016. Gender, Field, and Habitus: How Gendered Dispositions Reproduce Fields of Cultural Production. *Sociological Forum* **31**:2, 330-353. [Crossref]
68. Guillaume Beaulac, Tim Kenyon. 2016. The Scope of Debiasing in the Classroom. *Topoi* . [Crossref]
69. Christopher R Matthews. 2016. Exploring the pastiche hegemony of men. *Palgrave Communications* **2**, 16022. [Crossref]
70. Anthony Heyes, John A. List. 2016. Supply and Demand for Discrimination: Strategic Revelation of Own Characteristics in a Trust Game. *American Economic Review* **106**:5, 319-323. [Abstract] [View PDF article] [PDF with links]
71. Iris Bohnet, Alexandra van Geen, Max Bazerman. 2016. When Performance Trumps Gender Bias: Joint vs. Separate Evaluation. *Management Science* **62**:5, 1225-1234. [Crossref]
72. Ralph Hertwig, Christoph Engel. 2016. Homo Ignorans. *Perspectives on Psychological Science* **11**:3, 359-372. [Crossref]
73. Linda M. Dunn-Jensen, Scott Jensen, Mikelle A. Calhoun, Katherine C. Ryan. 2016. Revealing Gender Bias: An Experiential Exercise. *Organization Management Journal* **13**:2, 101-114. [Crossref]
74. Ana Rute Cardoso, Paulo Guimarães, Pedro Portugal. 2016. What drives the gender wage gap? A look at the role of firm and job-title heterogeneity. *Oxford Economic Papers* **68**:2, 506-524. [Crossref]
75. . Career Strategies for Women 36-43. [Crossref]
76. Luc Bovens. 2016. Selection under Uncertainty: Affirmative Action at Shortlisting Stage: Table 1. *Mind* **125**:498, 421-437. [Crossref]

77. Vikki Boliver. 2016. Exploring Ethnic Inequalities in Admission to Russell Group Universities. *Sociology* **50**:2, 247-266. [Crossref]

78. Alice H. Eagly. 2016. When Passionate Advocates Meet Research on Diversity, Does the Honest Broker Stand a Chance?. *Journal of Social Issues* **72**:1, 199-222. [Crossref]

79. Magali Fassiotto, Elizabeth Otto Hamel, Manwai Ku, Shelley Correll, Daisy Grewal, Philip Lavori, V.J. Periyakoil, Allan Reiss, Christy Sandborg, Gregory Walton, Marilyn Winkleby, Hannah Valantine. 2016. Women in Academic Medicine: Measuring Stereotype Threat Among Junior Faculty. *Journal of Women's Health* **25**:3, 292-298. [Crossref]

80. Sun-Young Park. 2016. A Gender and The Beauty Showed in Politician's Facial Appearance as an Element of Winning in the Election Process. *Journal of Digital Convergence* **14**:2, 397-414. [Crossref]

81. Robert Legg, David Jeffery. 2016. Suleika and Hatem revisited: uncovering the material advantages of identifying as a male composer. *Music Education Research* 1-10. [Crossref]

82. Jurjen J. A. Kamphorst, Otto H. Swank. 2016. Don't Demotivate, Discriminate. *American Economic Journal: Microeconomics* **8**:1, 140-165. [Abstract] [View PDF article] [PDF with links]

83. Diana L. Miller. 2016. Gender and the Artist Archetype: Understanding Gender Inequality in Artistic Careers. *Sociology Compass* **10**:2, 119-131. [Crossref]

84. Bentley Coffey, Patrick A. McLaughlin. 2016. The Effect on Lawyers Income of Gender Information Contained in First Names. *Review of Law & Economics*, ahead of print. [Crossref]

85. Elisabetta Addis. Economics, Gender, and Power 121-136. [Crossref]

86. Samuel V. Bruton. 2015. Looks-Based Hiring and Wrongful Discrimination. *Business and Society Review* **120**:4, 607-635. [Crossref]

87. Aidan Byrne, Alessandra Tanesini. 2015. Instilling new habits: addressing implicit bias in healthcare professionals. *Advances in Health Sciences Education* **20**:5, 1255-1262. [Crossref]

88. Kathy Lund Dean, Jeanie M. Forray. 2015. Harry Potter in the Academy. *Journal of Management Education* **39**:6, 675-680. [Crossref]

89. Bjorn Tyrefors Hinnerich, Erik Höglin, Magnus Johannesson. 2015. Discrimination against students with foreign backgrounds: evidence from grading in Swedish public high schools. *Education Economics* **23**:6, 660-676. [Crossref]

90. Thomas Breda, Son Thierry Ly. 2015. Professors in Core Science Fields Are Not Always Biased against Women: Evidence from France. *American Economic Journal: Applied Economics* **7**:4, 53-75. [Abstract] [View PDF article] [PDF with links]

91. Ian Ayres, Mahzarin Banaji, Christine Jolls. 2015. Race effects on eBay. *The RAND Journal of Economics* **46**:4, 891-917. [Crossref]

92. Maja Spener. 2015. Calibrating Introspection. *Philosophical Issues* **25**:1, 300-321. [Crossref]

93. Zheng Fang, Chris Sakellariou. 2015. Glass Ceilings versus Sticky Floors: Evidence from Southeast Asia and an International Update. *Asian Economic Journal* **29**:3, 215-242. [Crossref]

94. Andreas Leibbrandt, John A. List. 2015. Do Women Avoid Salary Negotiations? Evidence from a Large-Scale Natural Field Experiment. *Management Science* **61**:9, 2016-2024. [Crossref]

95. SIEW CHING GOY, GERAINT JOHNES. 2015. DIFFERENCES IN DECLINE: QUANTILE REGRESSION OF MALE–FEMALE EARNINGS DIFFERENTIAL IN MALAYSIA. *The Singapore Economic Review* **60**:04, 1550054. [Crossref]

96. Matthew A. Kraft. 2015. Teacher Layoffs, Teacher Quality, and Student Achievement: Evidence from a Discretionary Layoff Policy. *Education Finance and Policy* 1-41. [Crossref]

97. Lin Xiu, Gerui (Grace) Kang, Alan C. Roline. 2015. Who negotiates a higher starting salary?. *Nankai Business Review International* **6**:3, 240-255. [Crossref]

98. Luc Behaghel, Bruno Crépon, Thomas Le Barbanchon. 2015. Unintended Effects of Anonymous Résumés. *American Economic Journal: Applied Economics* **7**:3, 1-27. [Abstract] [View PDF article] [PDF with links]
99. Alex Bryson, Arnaud Chevalier. 2015. Is there a taste for racial discrimination amongst employers?. *Labour Economics* **34**, 51-63. [Crossref]
100. stephen bates, heather savigny. 2015. introduction: women in european political science. *European Political Science* **14**:2, 75-78. [Crossref]
101. Christina Scharff. 2015. Blowing your own Trumpet: Exploring the Gendered Dynamics of Self-Promotion in the Classical Music Profession. *The Sociological Review* **63**:1_suppl, 97-112. [Crossref]
102. Kenneth S. Rhee, Tracey H. Sigler. 2015. Untangling the relationship between gender and leadership. *Gender in Management: An International Journal* **30**:2, 109-134. [Crossref]
103. Shireen Kanji, Sandra Hupka-Brunner. 2015. Young women's strong preference for children and subsequent occupational gender segregation. *Equality, Diversity and Inclusion: An International Journal* **34**:2, 124-140. [Crossref]
104. E. S. Darling. 2015. Use of double-blind peer review to increase author diversity. *Conservation Biology* **29**:1, 297-299. [Crossref]
105. Darlene F. Saporu, Joan M. Herbers. 2015. What's in a Title? Gender Micro-Inequities in a University Human Resources Database. *NASPA Journal About Women in Higher Education* **8**:1, 101-116. [Crossref]
106. Kenneth Elpus. 2015. National estimates of male and female enrolment in American high school choirs, bands and orchestras. *Music Education Research* **17**:1, 88-102. [Crossref]
107. Sophie Ponthieux, Dominique Meurs. Gender Inequality 981-1146. [Crossref]
108. J. A. Flory, A. Leibbrandt, J. A. List. 2015. Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions. *The Review of Economic Studies* **82**:1, 122. [Crossref]
109. Douglas S. Massey. Inequality, Social 908-913. [Crossref]
110. Francine D. Blau. Gender, Economics of 757-763. [Crossref]
111. Jan W Peters, Nancy J Lane. 1. [Crossref]
112. Lekelia Danielle Jenkins. 2015. From conflict to collaboration: The role of expertise in fisheries management. *Ocean & Coastal Management* **103**, 123-133. [Crossref]
113. Maria De Paola, Vincenzo Scoppa. 2015. Gender Discrimination and Evaluators' Gender: Evidence from Italian Academia. *Economica* **82**:325, 162-188. [Crossref]
114. Diana L. Miller. 2014. Symbolic Capital and Gender: Evidence from Two Cultural Fields. *Cultural Sociology* **8**:4, 462-482. [Crossref]
115. Alastair Brown. 2014. Double-blind under review. *Nature Nanotechnology* **9**:11, 871-872. [Crossref]
116. Stijn Baert. 2014. Career lesbians. Getting hired for not having kids?. *Industrial Relations Journal* **45**:6, 543-561. [Crossref]
117. Andrea Romei, Salvatore Ruggieri. 2014. A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review* **29**:05, 582-638. [Crossref]
118. Joya Misra, Marta Murray-Close. 2014. The Gender Wage Gap in the United States and Cross Nationally. *Sociology Compass* **8**:11, 1281-1295. [Crossref]
119. Ghazala Azmat, Barbara Petrongolo. 2014. Gender and the labor market: What have we learned from field and lab experiments?. *Labour Economics* **30**, 32-40. [Crossref]
120. Geoffrey Lightfoot, Tomasz Piotr Wisniewski. 2014. Information asymmetry and power in a surveillance society. *Information and Organization* **24**:4, 214-235. [Crossref]

121. Nilanjana Dasgupta, Jane G. Stout. 2014. Girls and Women in Science, Technology, Engineering, and Mathematics. *Policy Insights from the Behavioral and Brain Sciences* **1**:1, 21-29. [Crossref]

122. Yuval Arbel, Yossef Tobol, Erez Siniver. 2014. Social involvement, level of income and employment among immigrants. *International Journal of Manpower* **35**:6, 798-816. [Crossref]

123. Roberto M. Fernandez, Roman V. Galperin. The Causal Status of Social Capital in Labor Markets 445-462. [Crossref]

124. HILLARY N. MORGAN, KURT W. ROTTHOFF. 2014. THE HARDER THE TASK, THE HIGHER THE SCORE: FINDINGS OF A DIFFICULTY BIAS. *Economic Inquiry* **52**:3, 1014-1026. [Crossref]

125. Katherine Hansen, Margaret Gerbasi, Alexander Todorov, Elliott Kruse, Emily Pronin. 2014. People Claim Objectivity After Knowingly Using Biased Strategies. *Personality and Social Psychology Bulletin* **40**:6, 691-699. [Crossref]

126. Chia-Jung Tsay. 2014. The vision heuristic: Judging music ensembles by sight alone. *Organizational Behavior and Human Decision Processes* **124**:1, 24-33. [Crossref]

127. Chiara Mussida, Matteo Picchio. 2014. The trend over time of the gender wage gap in Italy. *Empirical Economics* **46**:3, 1081-1110. [Crossref]

128. Lawrence M. Kahn. 2014. New Evidence on Gender and the Labor Market: A Symposium. *ILR Review* **67**:2, 283-286. [Crossref]

129. Balázs Kovács, Amanda J. Sharkey. 2014. The Paradox of Publicity. *Administrative Science Quarterly* **59**:1, 1-33. [Crossref]

130. Stéphane Robin, Agnieszka Rusinowska, Marie Claire Villeval. 2014. Ingratiation: Experimental evidence. *European Economic Review* **66**, 16-38. [Crossref]

131. Ming D. Leung. 2014. Dilettante or Renaissance Person? How the Order of Job Experiences Affects Hiring in an External Labor Market . *American Sociological Review* **79**:1, 136-158. [Crossref]

132. Thomas Biegert. 2014. On the outside looking in? Transitions out of non-employment in the United Kingdom and Germany. *Journal of European Social Policy* **24**:1, 3-18. [Crossref]

133. Nicole M. Lindner, Alexander Graser, Brian A. Nosek. 2014. Age-Based Hiring Discrimination as a Function of Equity Norms and Self-Perceived Objectivity. *PLoS ONE* **9**:1, e84752. [Crossref]

134. . References 197-203. [Crossref]

135. Y. Rubinstein, D. Brenner. 2014. Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labour Market Discrimination. *The Review of Economic Studies* **81**:1, 389-425. [Crossref]

136. Paola Profeta, Livia Amidani Aliberti, Alessandra Casarico, Marilisa D'Amico, Anna Puccio. Quotas on Boards: Evidence from the Literature 41-67. [Crossref]

137. Crystal L. Hoyt. Social Identities and Leadership: The Case of Gender 71-91. [Crossref]

138. Heather Savigny. 2014. Women, know your limits: cultural sexism in academia. *Gender and Education* **26**:7, 794. [Crossref]

139. Erin L. Cadwalader, Amanda C. Bryant-Friedrich. Improving Transparency and Equity in Scholarly Recognition by Scientific Societies 245-254. [Crossref]

140. Adrian Masters. 2013. Statistical discrimination from composition effects in the market for low-skilled workers. *Labour Economics* . [Crossref]

141. Philippe Jacquart, J. Scott Armstrong. 2013. The Ombudsman: Are Top Executives Paid Enough? An Evidence-Based Review. *Interfaces* **43**:6, 580-589. [Crossref]

142. Eric Yanfei Zhao, Masakazu Ishihara, Michael Lounsbury. 2013. Overcoming the Illegitimacy Discount: Cultural Entrepreneurship in the US Feature Film Industry. *Organization Studies* **34**:12, 1747-1776. [Crossref]

143. Kristen Monroe. 2013. Tougher Standards for Female Scholars? The Psychology Behind Them and Policies to Eliminate Them. *International Studies Perspectives* **14**:4, 476-484. [Crossref]

144. Jonathan Guryan, Kerwin Kofi Charles. 2013. Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots. *The Economic Journal* **123**:572, F417-F432. [Crossref]

145. Asaf Zussman. 2013. Ethnic Discrimination: Lessons from the Israeli Online Market for Used Cars. *The Economic Journal* **123**:572, F433-F468. [Crossref]

146. Silvia Knobloch-Westerwick, Carroll J. Glynn, Michael Huge. 2013. The Matilda Effect in Science Communication. *Science Communication* **35**:5, 603-625. [Crossref]

147. Chia-Jung Tsay. 2013. Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences* **110**:36, 14580-14585. [Crossref]

148. C. Bram Cadsby, Maroš Servátka, Fei Song. 2013. How competitive are female professionals? A tale of identity conflict. *Journal of Economic Behavior & Organization* **92**, 284-303. [Crossref]

149. Jevin D. West, Jennifer Jacquet, Molly M. King, Shelley J. Correll, Carl T. Bergstrom. 2013. The Role of Gender in Scholarly Authorship. *PLoS ONE* **8**:7, e66212. [Crossref]

150. David A. Matsa,, Amalia R. Miller. 2013. A Female Style in Corporate Leadership? Evidence from Quotas. *American Economic Journal: Applied Economics* **5**:3, 136-169. [Abstract] [View PDF article] [PDF with links]

151. Roxana Barbulescu, Matthew Bidwell. 2013. Do Women Choose Different Jobs from Men? Mechanisms of Application Segregation in the Market for Managerial Workers. *Organization Science* **24**:3, 737-756. [Crossref]

152. Giacomo Negro, Ming D. Leung. 2013. "Actual" and Perceptual Effects of Category Spanning. *Organization Science* **24**:3, 684-696. [Crossref]

153. Victor Lavy. 2013. Gender Differences in Market Competitiveness in a Real Workplace: Evidence from Performance-based Pay Tournaments among Teachers*. *The Economic Journal* **123**:569, 540-573. [Crossref]

154. Amine Ouazad, Lionel Page. 2013. Students' Perceptions of Teacher Biases: Experimental Economics in Schools. *Journal of Public Economics* . [Crossref]

155. Calvin K. Lai, Kelly M. Hoffman, Brian A. Nosek. 2013. Reducing Implicit Prejudice. *Social and Personality Psychology Compass* **7**:5, 315-330. [Crossref]

156. Manuel Bagues, Maria J. Perez-Villadoniga. 2013. Why do I like people like me?. *Journal of Economic Theory* **148**:3, 1292-1299. [Crossref]

157. Robert O. Deaner. 2013. Distance Running as an Ideal Domain for Showing a Sex Difference in Competitiveness. *Archives of Sexual Behavior* **42**:3, 413-428. [Crossref]

158. Eric P. Baumer. 2013. Reassessing and Redirecting Research on Race and Sentencing. *Justice Quarterly* **30**:2, 231-261. [Crossref]

159. Feng Li, Venky Nagar. 2013. Diversity and Performance. *Management Science* **59**:3, 529-544. [Crossref]

160. Marco Castillo, Ragan Petrie, Maximo Torero, Lise Vesterlund. 2013. Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics* **99**, 35-48. [Crossref]

161. Oscar Afonso. 2013. SCALE-INDEPENDENT TECHNOLOGICAL-KNOWLEDGE BIAS, HUMAN-CAPITAL ACCUMULATION AND GENDER INEQUALITY. *Metroeconomica* **64**:1, 125-151. [Crossref]

162. NABANITA DATTA GUPTA, ANDERS POULSEN, MARIE CLAIRE VILLEVAL. 2013. GENDER MATCHING AND COMPETITIVENESS: EXPERIMENTAL EVIDENCE. *Economic Inquiry* **51**:1, 816-835. [Crossref]

163. Catherine P. Mulder. 2013. Unions' Resistance to Capital and the Potential for Class Transformation. *Rethinking Marxism* **25**:1, 114-120. [Crossref]

164. Muriel Niederle, Carmit Segal, Lise Vesterlund. 2013. How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. *Management Science* **59**:1, 1-16. [Crossref]

165. K. Inman, N. Rudin. Sequential Unmasking: Minimizing Observer Effects in Forensic Science 542-548. [Crossref]

166. Nicolas Jacquemet, Constantine Yannelis. 2012. Indiscriminate discrimination: A correspondence test for ethnic homophily in the Chicago labor market. *Labour Economics* **19**:6, 824-832. [Crossref]

167. Rema N. Hanna,, Leigh L. Linden. 2012. Discrimination in Grading. *American Economic Journal: Economic Policy* **4**:4, 146-168. [Abstract] [View PDF article] [PDF with links]

168. Marie-Pierre Dargnies. 2012. Men Too Sometimes Shy Away from Competition: The Case of Team Competition. *Management Science* **58**:11, 1982-2000. [Crossref]

169. Benjamin Bolitzer, Erin M. Godtland. 2012. Understanding the Gender–Pay Gap in the Federal Workforce Over the Past 20 Years. *The American Review of Public Administration* **42**:6, 730-746. [Crossref]

170. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, J. Handelsman. 2012. Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences* **109**:41, 16474-16479. [Crossref]

171. Stefan Eriksson, Jonas Lagerström. 2012. Detecting discrimination in the hiring process: evidence from an Internet-based search channel. *Empirical Economics* **43**:2, 537-563. [Crossref]

172. Tessa V. West, Madeline E. Heilman, Lindy Gullett, Corinne A. Moss-Racusin, Joe C. Magee. 2012. Building blocks of bias: Gender composition predicts male and female group members' evaluations of each other and the group. *Journal of Experimental Social Psychology* **48**:5, 1209-1212. [Crossref]

173. Juan J. Dolado, Florentino Felgueroso, Miguel Almunia. 2012. Are men and women-economists evenly distributed across research fields? Some new empirical evidence. *SERIEs* **3**:3, 367-393. [Crossref]

174. Sandra A. Howard. 2012. The Effect of Selected Nonmusical Factors on Adjudicators' Ratings of High School Solo Vocal Performances. *Journal of Research in Music Education* **60**:2, 166-185. [Crossref]

175. Lisa D. Cook. 2012. Overcoming Discrimination by Consumers during the Age of Segregation: The Example of Garrett Morgan. *Business History Review* **86**:02, 211-234. [Crossref]

176. Marc Bendick, Ana P. Nunes. 2012. Developing the Research Basis for Controlling Bias in Hiring. *Journal of Social Issues* **68**:2, 238-262. [Crossref]

177. Nilanjana Dasgupta, Jane G. Stout. 2012. Contemporary Discrimination in the Lab and Field: Benefits and Obstacles of Full-Cycle Social Psychology. *Journal of Social Issues* **68**:2, 399-412. [Crossref]

178. Petra Moser. 2012. Taste-based discrimination evidence from a shift in ethnic preferences after WWI. *Explorations in Economic History* **49**:2, 167-188. [Crossref]

179. OLGA ALONSO-VILLAR, CORAL DEL RIO, CARLOS GRADIN. 2012. The Extent of Occupational Segregation in the United States: Differences by Race, Ethnicity, and Gender. *Industrial Relations: A Journal of Economy and Society* **51**:2, 179-212. [Crossref]

180. Randall Akee, Mutlu Yuksel. 2012. The Decreasing Effect of Skin Tone on Women's Full-Time Employment. *ILR Review* **65**:2, 398-426. [Crossref]

181. Anne E. Lincoln, Stephanie Pincus, Janet Bandows Koster, Phoebe S. Leboy. 2012. The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science* **42**:2, 307-320. [Crossref]

182. Brian A. Nosek, Rachel G. Riskind. 2012. Policy Implications of Implicit Social Cognition. *Social Issues and Policy Review* **6**:1, 113-147. [Crossref]

183. Leng Lee. 2012. Decomposing wage differentials between migrant workers and urban workers in urban China's labor markets. *China Economic Review* . [Crossref]

184. Valerae Olive Lewis, Susan A Scherl, Mary I O'Connor. 2012. Women in Orthopaedics-Way Behind the Number Curve. *The Journal of Bone and Joint Surgery-American Volume* **94**:5, e30-1-7. [Crossref]

185. L. Balafoutas, M. Sutter. 2012. Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory. *Science* **335**:6068, 579-582. [Crossref]

186. Thomas Cornelissen, Uwe Jirjahn. 2012. September 11th and the earnings of Muslims in Germany —The moderating role of education and firm size. *Journal of Economic Behavior & Organization* **81**:2, 490-504. [Crossref]

187. Hannah Riley Bowles. 2012. Claiming authority: How women explain their ascent to top business leadership positions. *Research in Organizational Behavior* **32**, 189-212. [Crossref]

188. Fredrik Carlsson, Åsa Löfgren, Thomas Sterner. 2012. Discrimination in Scientific Review: A Natural Field Experiment on Blind versus Non-Blind Reviews*. *The Scandinavian Journal of Economics* no-no. [Crossref]

189. Richard F. Martell, Cynthia G. Emrich, James Robison-Cox. 2012. From bias to exclusion: A multilevel emergent theory of gender segregation in organizations. *Research in Organizational Behavior* **32**, 137-162. [Crossref]

190. 2012. Pathways into a Gendered Occupation. *International Journal of Social and Organizational Dynamics in IT* **2**:4, 34-51. [Crossref]

191. Olof Åslund, Oskar Nordströum Skans. 2012. Do Anonymous Job Application Procedures Level the Playing Field?. *ILR Review* **65**:1, 82-107. [Crossref]

192. Wei Zhao, Xueguang Zhou. 2011. Status Inconsistency and Product Valuation in the California Wine Market. *Organization Science* **22**:6, 1435-1448. [Crossref]

193. Donna Bobbitt-Zeher. 2011. Gender Discrimination at Work. *Gender & Society* **25**:6, 764-786. [Crossref]

194. Philip Oreopoulos. 2011. Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes. *American Economic Journal: Economic Policy* **3**:4, 148-171. [Abstract] [View PDF article] [PDF with links]

195. Amy J.C. Cuddy, Peter Glick, Anna Beninger. 2011. The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior* . [Crossref]

196. Jesper B. Sørensen, Amanda J. Sharkey. 2011. The Perils of False Certainty. *Sociological Methods & Research* **40**:4, 635-645. [Crossref]

197. Shamena Anwar. 2011. Testing for Discrimination: Evidence from the Game Show Street Smarts. *Journal of Economic Behavior & Organization* . [Crossref]

198. Kerwin Kofi Charles, Jonathan Guryan. 2011. Studying Discrimination: Fundamental Challenges and Recent Progress. *Annual Review of Economics* **3**:1, 479-511. [Crossref]

199. Muriel Niederle, Lise Vesterlund. 2011. Gender and Competition. *Annual Review of Economics* **3**:1, 601-630. [Crossref]

200. Maria Charles. 2011. A World of Difference: International Trends in Women's Economic Status. *Annual Review of Sociology* **37**:1, 355-371. [Crossref]

201. Martín Moreno, Hugo Ñopo, Jaime Saavedra, Máximo Torero. 2011. Detecting Gender and Racial Discrimination in Hiring through Monitoring Intermediation Services: The Case of Selected Occupations in Metropolitan Lima, Peru. *World Development* . [Crossref]

202. D. James Greiner, Donald B. Rubin. 2011. Causal Effects of Perceived Immutable Characteristics. *Review of Economics and Statistics* **93**:3, 775-785. [Crossref]

203. David M. Waguespack, Olav Sorenson. 2011. The Ratings Game: Asymmetry in Classification. *Organization Science* **22**:3, 541-553. [Crossref]

204. Sally Chapman, Felicia F. Dixon, Natalie Foster, Valerie J. Kuck, Deborah A. McCarthy, Nancy M. Tooney, Janine P. Buckner, Susan A. Nolan, Cecilia H. Marzabadi. 2011. Female Faculty Members in University Chemistry Departments: Observations and Conclusions Based on Site Visits. *Journal of Chemical Education* **88**:6, 716-720. [Crossref]

205. SUBHASISH DUGAR, HAIMANTI BHATTACHARYA, DAVID REILEY. 2011. CAN'T BUY ME LOVE? A FIELD EXPERIMENT EXPLORING THE TRADE-OFF BETWEEN INCOME AND CASTE-STATUS IN AN INDIAN MATRIMONIAL MARKET. *Economic Inquiry* no-no. [Crossref]

206. C. R. Taylor, H. Yildirim. 2011. Subjective Performance and the Value of Blind Evaluation. *The Review of Economic Studies* **78**:2, 762-794. [Crossref]

207. Chia-Jung Tsay, Mahzarin R. Banaji. 2011. Naturals and strivers: Preferences and beliefs about sources of achievement. *Journal of Experimental Social Psychology* **47**:2, 460-465. [Crossref]

208. Leo Kaas, Christian Manger. 2011. Ethnic Discrimination in Germany's Labour Market: A Field Experiment. *German Economic Review* no-no. [Crossref]

209. S. J. Ceci, W. M. Williams. 2011. Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences* **108**:8, 3157-3162. [Crossref]

210. Timothy S. Simcoe, Dave M. Waguespack. 2011. Status, Quality, and Attention: What's in a (Missing) Name?. *Management Science* **57**:2, 274-290. [Crossref]

211. Yang Yang, Alison M. Konrad. 2011. Understanding Diversity Management Practices: Implications of Institutional Theory and Resource-Based Theory. *Group & Organization Management* **36**:1, 6-38. [Crossref]

212. Barbara F. Reskin, Michelle L. Maroto. 2011. What Trends? Whose Choices?. *Gender & Society* **25**:1, 81-87. [Crossref]

213. John A. List, Imran Rasul. Field Experiments in Labor Economics 103-228. [Crossref]

214. Timo Gnambs, Bernad Batinic. 2011. Evaluation of measurement precision with Rasch-type models: The case of the short Generalized Opinion Leadership Scale. *Personality and Individual Differences* **50**:1, 53-58. [Crossref]

215. Barbara Bonnekessen. 2011. The NSF Career-Life Balance Initiative: A Critical Examination. *Journal of Workplace Rights* **16**:3, 299-313. [Crossref]

216. Mary Gatta. In the 'Blink' of an Eye — American High-End Small Retail Businesses and the Public Workforce System 49-67. [Crossref]

217. Benjamin L. Solow, John L. Solow, Todd B. Walker. 2010. Moving on up: The Rooney rule and minority hiring in the NFL#. *Labour Economics* . [Crossref]

218. Vera Borges. 2010. Trabalho, género, idade e arte: estudos empíricos sobre o teatro e a dança. *e-cadernos CES* :10. . [Crossref]

219. MANUEL F. BAGUES, BERTA ESTEVE-VOLART. 2010. Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment. *Review of Economic Studies* **77**:4, 1301-1328. [Crossref]

220. Lars Bo Jeppesen, Karim R. Lakhani. 2010. Marginality and Problem-Solving Effectiveness in Broadcast Search. *Organization Science* **21**:5, 1016-1033. [Crossref]

221. Guy Michaels,, Xiaojia Zhi. 2010. Freedom Fries. *American Economic Journal: Applied Economics* **2**:3, 256-281. [Abstract] [View PDF article] [PDF with links]

222. Crystal L. Hoyt. 2010. Women, Men, and Leadership: Exploring the Gender Gap at the Top. *Social and Personality Psychology Compass* **4**:7, 484-498. [Crossref]

223. Diana Bilimoria, Kimberly Buch. 2010. The Search is On: Engendering Faculty Diversity Through More Effective Search and Recruitment. *Change: The Magazine of Higher Learning* **42**:4, 27-32. [Crossref]

224. Markus Gangl. 2010. Causal Inference in Sociological Research. *Annual Review of Sociology* **36**:1, 21-47. [Crossref]

225. PAOLA SAPIENZA. 2010. Discussion of Self-Selection and the Forecasting Abilities of Female Equity Analysts. *Journal of Accounting Research* **48**:2, 437-443. [Crossref]

226. Destiny Peery, Jennifer Richeson. 2010. Broadening Horizons: Considerations for Creating a More Complete Science of Diversity. *Psychological Inquiry* **21**:2, 146-152. [Crossref]

227. Arthur Eisenkraft. 2010. Millikan Lecture 2009: Physics for all: From special needs to Olympiads. *American Journal of Physics* **78**:4, 328-337. [Crossref]

228. Vincent Yzerbyt, Stéphanie Demoulin. . [Crossref]

229. Deborah M. Kolb. 2009. Too Bad for the Women or Does It Have to Be? Gender and Negotiation Research over the Past Twenty-Five Years. *Negotiation Journal* **25**:4, 515-531. [Crossref]

230. Nicola Persico. 2009. Racial Profiling? Detecting Bias Using Statistical Evidence. *Annual Review of Economics* **1**:1, 229-254. [Crossref]

231. K. Anders Ericsson, Kiruthiga Nandagopal, Roy W. Roring. 2009. Toward a Science of Exceptional Achievement. *Annals of the New York Academy of Sciences* **1172**:1, 199-217. [Crossref]

232. MARLENE KIM. 2009. Race and Gender Differences in the Earnings of Black Workers. *Industrial Relations: A Journal of Economy and Society* **48**:3, 466-488. [Crossref]

233. Susan Sturm. 2009. Negotiating Workplace Equality: A Systemic Approach. *Negotiation and Conflict Management Research* **2**:1, 92-106. [Crossref]

234. Tim Dare, Justine Kingsbury. 2008. Putting the Burden of Proof in Its Place: When Are Differential Allocations Legitimate?. *The Southern Journal of Philosophy* **46**:4, 503-518. [Crossref]

235. Robert L. Nelson, Ellen C. Berrey, Laura Beth Nielsen. 2008. Divergent Paths: Conflicting Conceptions of Employment Discrimination in Law and the Social Sciences. *Annual Review of Law and Social Science* **4**:1, 103-122. [Crossref]

236. Madeline E. Heilman, Alice H. Eagly. 2008. Gender Stereotypes Are Alive, Well, and Busy Producing Workplace Discrimination. *Industrial and Organizational Psychology* **1**:04, 393-398. [Crossref]

237. RAYMOND PATERNOSTER, ROBERT BRAME. 2008. REASSESSING RACE DISPARITIES IN MARYLAND CAPITAL CASES*. *Criminology* **46**:4, 971-1008. [Crossref]

238. C. Elizabeth Hirsh, Youngjoo Cha. 2008. Understanding Employment Discrimination: A Multilevel Approach. *Sociology Compass* **2**:6, 1989-2007. [Crossref]

239. Muriel Niederle, Lise Vesterlund. 2008. Gender Differences in Competition. *Negotiation Journal* **24**:4, 447-463. [Crossref]

240. Tanya S. Rosenblat. 2008. The Beauty Premium: Physical Attractiveness and Gender in Dictator Games. *Negotiation Journal* **24**:4, 465-481. [Crossref]

241. Fiona Greig. 2008. Propensity to Negotiate and Career Advancement: Evidence from an Investment Bank that Women Are on a "Slow Elevator". *Negotiation Journal* **24**:4, 495-508. [Crossref]

242. Claudia S. Morrissey, Mary Lou Schmidt. 2008. Fixing the System, Not the Women: An Innovative Approach to Faculty Advancement. *Journal of Women's Health* **17**:8, 1399-1408. [Crossref]

243. Catherine Eckel, Angela C. M. de Oliveira, Philip J. Grossman. 2008. Gender and Negotiation in the Small: Are Women (Perceived to Be) More Cooperative than Men?. *Negotiation Journal* **24**:4, 429-445. [Crossref]

244. Melissa J. Williams, Gerald A. Mendelsohn. 2008. Gender Clues and Cues: Online Interactions as Windows into Lay Theories about Men and Women. *Basic and Applied Social Psychology* **30**:3, 278-294. [Crossref]

245. David Bjerk. 2008. Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Hiring and Promotion*. *The Economic Journal* **118**:530, 961-982. [Crossref]

246. THOMAS J. DOHMEN. 2008. THE INFLUENCE OF SOCIAL FORCES: EVIDENCE FROM THE BEHAVIOR OF FOOTBALL REFEREES. *Economic Inquiry* **46**:3, 411-424. [Crossref]

247. SALLY HASLANGER. 2008. Changing the Ideology and Culture of Philosophy: Not by Reason (Alone). *Hypatia* **23**:2, 210-223. [Crossref]

248. Sally Haslanger. 2008. Changing the Ideology and Culture of Philosophy: Not by Reason (Alone). *Hypatia: A Journal of Feminist Philosophy* **23**:2, 210-223. [Crossref]

249. David H. Autor, David Scarborough. 2008. Does Job Testing Harm Minority Workers? Evidence from Retail Establishments *. *Quarterly Journal of Economics* **123**:1, 219-277. [Crossref]

250. Stephanie Riegg Cellini, Signe-Mary McKernan, Caroline Ratcliffe. 2008. The dynamics of poverty in the United States: A review of data, methods, and findings. *Journal of Policy Analysis and Management* **27**:3, 577-605. [Crossref]

251. A JONES, K FRICK. 2008. Gender Bias in Economic Evaluation MethodsTime Costs and Productivity Loss. *Women's Health Issues* **18**:1, 1-3. [Crossref]

252. John J. Donohue. Anti-Discrimination Law 1-10. [Crossref]

253. Sheryl Ball. Status and Economics 1-6. [Crossref]

254. Francine D. Blau, Lawrence M. Kahn. Women's Work and Wages 1-14. [Crossref]

255. Arnaud Chevalier. 2007. Education, Occupation and Career Expectations: Determinants of the Gender Pay Gap for UK Graduates. *Oxford Bulletin of Economics and Statistics* **69**:6, 819-842. [Crossref]

256. Annick Vignes, Olivier Gergaud. 2007. Twilight of the Idols in the Market for Champagne: Dissonance or Consonance in Consumer Preferences?. *Journal of Wine Research* **18**:3, 147-162. [Crossref]

257. Judith L. Komaki. 2007. Daring to Dream: Promoting Social and Economic Justice at Work. *Applied Psychology* **56**:4, 624-662. [Crossref]

258. Muriel Niederle, Lise Vesterlund. 2007. Do Women Shy Away from Competition? Do Men Compete Too Much?*. *Quarterly Journal of Economics* **122**:3, 1067-1101. [Crossref]

259. K. Anders Ericsson, Roy W. Roring, Kiruthiga Nandagopal. 2007. Giftedness and evidence for reproducibly superior performance: an account based on the expert performance framework. *High Ability Studies* **18**:1, 3-56. [Crossref]

260. John J. Donohue. Chapter 18 Antidiscrimination Law 1387-1472. [Crossref]

261. Catherine Loughlin, Kara A. Arnold. 2007. Seeking the best: Leadership lessons from the military. *Human Resource Management* **46**:1, 147-167. [Crossref]

262. Anette Fasang. 2006. Recruitment in symphony orchestras: testing a gender neutral recruitment process. *Work, employment and society* **20**:4, 801-809. [Crossref]
263. Niki T. Dickerson. 2006. "WE ARE A FORCE TO BE RECKONED WITH": BLACK AND LATINA WOMEN'S LEADERSHIP IN THE CONTEMPORARY U.S. LABOR MOVEMENT. *WorkingUSA* **9**:3, 293-313. [Crossref]
264. Evalyn Gates. 2006. A scientific point of view. *Physics Today* **59**:4, 64-65. [Crossref]
265. Ingrid Robeyns. 2006. Three models of education. *School Field* **4**:1, 69-84. [Crossref]
266. Markus M. Mobius, Tanya S. Rosenblat. 2006. Why Beauty Matters. *American Economic Review* **96**:1, 222-235. [Abstract] [View PDF article] [PDF with links]
267. Morley Gunderson. 2006. Viewpoint: Male-female wage differentials: how can that be?. *Canadian Journal of Economics/Revue canadienne d'<html_ent glyph="@eacute;" ascii="e"/>conomique* **39**:1, 1-21. [Crossref]
268. Neil O. Alper, Gregory H. Wassall. Chapter 23 Artists' Careers and Their Labor Markets 813-864. [Crossref]
269. TROND PETERSEN, ISHAK SAPORTA, MARC-DAVID L. SEIDEL. 2005. Getting Hired: Sex and Race. *Industrial Relations* **44**:3, 416-443. [Crossref]
270. Luis Garicano, Ignacio Palacios-Huerta, Canice Prendergast. 2005. Favoritism Under Social Pressure. *Review of Economics and Statistics* **87**:2, 208-216. [Crossref]
271. Esther Duflo. 2005. Why Political Reservations?. *Journal of the European Economic Association* **3**:2-3, 668-678. [Crossref]
272. Barbara F. Reskin, Denise D. Bielby. 2005. A Sociological Perspective on Gender and Career Outcomes. *Journal of Economic Perspectives* **19**:1, 71-86. [Abstract] [View PDF article] [PDF with links]
273. Robin E. Bell, Jennifer D. Laird. 2005. Women, Work, and the Academy. *Eos, Transactions American Geophysical Union* **86**:30. . [Crossref]
274. Marianne Bertrand, Sendhil Mullainathan. 2004. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review* **94**:4, 991-1013. [Abstract] [View PDF article] [PDF with links]
275. Uri Gneezy, Aldo Rustichini. 2004. Gender and Competition at a Young Age. *American Economic Review* **94**:2, 377-381. [Citation] [View PDF article] [PDF with links]
276. John A. List. 2004. The Nature and Extent of Discrimination in the Marketplace: Evidence From the Field*. *Quarterly Journal of Economics* **119**:1, 49-89. [Crossref]
277. Trond Petersen, Ishak Saporta. 2004. The Opportunity Structure for Discrimination. *American Journal of Sociology* **109**:4, 852-901. [Crossref]
278. Judea Pearl. 2003. Statistics and causal inference: A review. *Test* **12**:2, 281-345. [Crossref]
279. David Neumark. 2003. Age Discrimination Legislation in the United States. *Contemporary Economic Policy* **21**:3, 297-317. [Crossref]
280. Kaivan Munshi. 2003. Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market*. *Quarterly Journal of Economics* **118**:2, 549-599. [Crossref]
281. W. Bentley Macleod. 2003. Optimal Contracting with Subjective Evaluation. *American Economic Review* **93**:1, 216-240. [Abstract] [View PDF article] [PDF with links]
282. Victor A. Ginsburgh, Jan C. van Ours. 2003. Expert Opinion and Compensation: Evidence from a Musical Competition. *American Economic Review* **93**:1, 289-296. [Citation] [View PDF article] [PDF with links]

283. P. A. Riach, J. Rich. 2002. Field Experiments of Discrimination in the Market Place*. *The Economic Journal* **112**:483, F480-F518. [Crossref]

284. Erik D. Craft. 2002. THE DEMAND FOR VANITY (PLATES): ELASTICITIES, NET REVENUE MAXIMIZATION, AND DEADWEIGHT LOSS. *Contemporary Economic Policy* **20**:2, 133-144. [Crossref]

285. F.D. Blau. Gender, Economics of 5995-6002. [Crossref]

286. Francine D. Blau,, Lawrence M. Kahn,. 2000. Gender Differences in Pay. *Journal of Economic Perspectives* **14**:4, 75-100. [Abstract] [View PDF article] [PDF with links]

287. Trond Petersen, Ishak Saporta, Marc-David L. Seidel. 2000. Offering a Job: Meritocracy and Social Networks. *American Journal of Sociology* **106**:3, 763-816. [Crossref]

288. Brenda Parker. Beyond the class act: Gender and race in the 'creative city' discourse 201-232. [Crossref]

289. Denise Lewin Loyd, Katherine W. Phillips. Managing Perceptions of Ethical Behavior in Evaluative Groups: The Implications for Diversity in Organizations 225-245. [Crossref]

International Journal of Obesity (2001) 25, 1246–1252
 0307–0565/01 $15.00
www.nature.com/ijo

PAPER

Weighing the care: physicians' reactions to the size of a patient

MR Hebl[1]* and J Xu[2]

[1]*Rice University, Houston, Texas, USA; and* [2]*University of Texas, Houston School of Public Health, Houston, Texas, USA*

OBJECTIVE: To examine how the weight of a patient affects both the attitudes that physicians hold as well as the treatments that they intend to prescribe.
DESIGN: In a six-cell randomized design, physicians evaluated a medical chart of a male or female patient, depicted as either average weight, overweight or obese, who presented with a migraine headache.
SUBJECTS: A total of 122 physicians affiliated with one of three hospitals located in the Texas Medical Center of Houston completed the experiment.
MEASUREMENTS: Using a standard medical procedure form, physicians indicated how long they would spend with the patient and which of 41 medical tests and procedures they would conduct. They also indicated their affective and behavioral reactions to the patient.
RESULTS: The weight of a patient significantly affected how physicians viewed and treated them. Although physicians prescribed more tests for heavier patients, $F(2, 107) = 3.65$, $P < 0.03$, they simultaneously indicated that they would spend less time with them, $F(2, 107) = 8.38$, $P < 0.001$, and viewed them significantly more negatively on 12 of the 13 indices.
CONCLUSION: This study reveals that physicians continue to play an influential role in lowering the quality of healthcare that overweight and obese patients receive. As the girth of America continues to increase, continued research and improvements in the quality of such healthcare deserve attention.
International Journal of Obesity (2001) **25,** 1246–1252

Keywords: physician attitudes; prejudice; care; stigma

Introduction

Most current reports suggest that almost one in every two American adults is overweight (BMI > 25.0 kg/m^2)[1] and one in every five adults is obese (BMI > 30 kg/m^2).[2] The list of very serious medical conditions to which obesity has been linked is extensive and recent reports even link obesity with death.[3–5] The social repercussions impose additional risks to the well being of those who are overweight.[6–8] Deflections from thin societal standards trigger attributions of weak character, an undisciplined nature and laziness.[9,10] As a whole, then, overweight individuals are viewed as having a physical, emotional and moral impairment, and are discriminated against in diverse domains, including employment,[11,12] education[13] and personal relationships.[14]

The current research attempts to clarify further this solemn picture by specifically examining how physicians respond to overweight and obese patients. While the general public's views on obesity are important in understanding the lives and experiences of those who are overweight, we believe it is critical to examine how physicians view such patients and, more importantly, how they propose to treat such patients. On one hand, physicians should be immune to the stigma of obesity — they encounter it frequently in their patients, they have access to recent empirical studies that suggest genetic influences and some uncontrollable elements of obesity,[15–17] and they are trained to treat patients warmly and professionally. On the other hand, physicians may be unaffected by or reactive against any suggestion or publication implying that they should accept obesity in their patients. The current research addresses physicians' reactions to patients of varying weights, and

*Correspondence: MR Hebl, Rice University, Department of Psychology, 6100 S. Main Street, MS 25, Houston, TX 77005, USA.
E-mail: hebl@rice.edu
Received 21 June 2000; revised 13 December 2000; accepted 21 February 2001

assesses whether the quality of healthcare delivery is uniform for patients of varying sizes.

Preliminary evidence collected more than 10 y ago suggests that health professionals hold negative attitudes toward those who are overweight.[18–22] Such studies show that nurses, medical students and physicians all hold negative attitudes toward and stereotypes regarding obese patients. The current study extends past research in two ways. First, it remains unclear how these attitudes toward heavier patients translate into physicians' intended behaviors and medical care delivery. Therefore, the current study examines the perceptions of and recommended care for patients of varying weights. Second, past studies do not differentiate between the attitudes that physicians hold toward their overweight and obese patients, so this study addresses physician reactions to such variations in size. In particular, we manipulate the weight of ostensible patients to be of average weight (BMI = 23), overweight (BMI = 30), and obese (BMI = 36). Examining multiple levels of weight enables us to determine if weight acts in a predicted linear fashion such that, the heavier the patient is, the more physicians perceive and respond to them negatively.

Methods

Participants

Participants were 122 primary care physicians, who were affiliated with one of three large hospitals located in the Texas Medical Center of Houston. These participants ranged in age (29 to 76-y-old, $M = 45.86$, s.d. = 9.98), years of practice (0–46 y, $M = 14.64$, s.d. = 10.34), and gender (90 males, 30 females, two unidentified). Physicians' informed consents were obtained by having them return the experimental packet, and all procedures were approved by the institutional review board at the University of Texas-School of Medicine.

Procedure

Participants were randomly assigned to receive by mail an experimental packet containing one of six variations of medical forms. Seventeen physicians reviewed the file of an average-weight female patient (1.65 m, 61.3 kg), 17 reviewed an average weight male patient (1.78 m, 72.6 kg), 21 reviewed an overweight female patient (1.65 m, 79.5 kg), 19 reviewed an overweight male patient (1.78 m, 95.3 kg), 29 reviewed an obese female patient (1.65 m, 95.3 kg), and 19 reviewed an obese male patient (1.78 m, 113.5 kg). A chi-square conducted on these response rates did not reveal a significant difference across conditions in response rate, $\chi^2 = 1.01$, $P = 0.60$.

All physicians were told the study involved responses to the completion of a medical chart that physicians characteristically review prior to meeting with a patient. Hence, physicians were sent a standard medical chart form presenting a patient and describing his/her age, weight, height, gender, race, marital status, blood pressure, temperature, pulse, medical history and family health background. The patient was depicted as healthy with two exceptions. First, two-thirds of the patients were depicted as overweight or obese, and second, the patients' presenting condition was disclosed, indicating a medical history of two migraine headaches spread over a time period of 2 y. This medical condition was chosen because it was not considered to be linked with body weight by the consulting physicians that we asked both prior to and during pretesting. Physicians were asked to recommend medical procedures for the patient and indicate their attitudes and plans regarding the patients, whom they believed was authentic.

Materials

Physicians completed two forms, the 'Medical procedures form' and the 'Patient follow-up questionnaire'. The Medical procedures form asked physicians to indicate, by checking boxes, all of the tests, procedures and referrals that they planned to conduct in caring for the patient. Forty-one tests and medical procedures were listed and ranged in the degree to which they were relevant to migraine headaches and weight-related conditions (see Table 1 for a complete list). The particular type and overall number of tests recommended by the physician were assessed.

The Patient follow-up questionnaire first asked physicians to indicate the amount of time they would spend with the patient. Then it assessed 13 affective and behavioral reactions that the physician had toward the patient. Physicians were asked: (a) to judge the health of the patient; (b) how well the patient took care of himself/herself; (c) how self-disciplined they perceived the patient to be; (d) the extent to which the physician would have to be strict; (e) the seriousness of the medical problem; (f) the extent to which they thought seeing this patient would be a waste of their time; (g) the extent to which seeing such patients would result in affinity for their job; (h) the level of patience that they would have for the patient; (i) the extent to which the patient would be annoying; (j) how much personal desire they had to help the patient; (k) the likelihood with which the patient would comply with medical advice; (l) whether the patient would benefit from psychological counseling; and (m) the overall level of positivity toward the patient. In answering all of these questions, physicians responded on nine-point scales that were anchored by (1) = 'Not at all', (5) = 'Some – Somewhat' and (9) = 'Extremely'.

Results

To examine responses on the Medical procedures form, we summed the overall number of procedures that each physician recommended and conducted a 2 (gender) × 3 (weight) ANOVA on the number of tests run. Across all of the analyses conducted, only one gender difference emerged, indicating that physicians recommended running more tests and procedures on women ($M = 11.76$) than on men ($M = 10.06$),

Table 1 Medical procedures physicians recommended for patient and chi-square results

	Average, BMI = 23, n = *34* n; *(%)*	*Overweight, BMI = 30,* n = *40* n; *(%)*	*Obese, BMI = 36,* n = *48* n; *(%)*	*Pearson chi-square (with 2 d.f.),* χ^2; P
1. Cholesterol level	3 (9)	14 (35)	16 (33)	7.96; $P = 0.02$
2. Tryglycerides level	3 (9)	14 (35)	15 (31)	7.54; $P = 0.02$
3. Body fat percentage	1 (3)	8 (25)	9 (19)	5.26; $P = 0.07$
4. Glucose level	11 (32)	19 (48)	21 (44)	1.86; $P = 0.40$
5. Dietary intake info	21 (62)	26 (65)	38 (79)	3.47; $P = 0.18$
6. Stress assessment	21 (62)	28 (70)	36 (75)	1.65; $P = 0.44$
7. Prescription of anti-depressants	2 (6)	7 (18)	10 (26)	3.55; $P = 0.17$
8. Consult about weight loss	1 (3)	14 (35)	20 (42)	15.79; $P < 0.01$
9. Consult about exercise	2 (6)	9 (23)	13 (27)	5.96; $P = 0.05$
10. Consult about nutrition	1 (3)	12 (30)	15 (31)	10.69; $P < 0.01$
11. Refer to psychologist	1 (3)	6 (15)	11 (23)	6.32; $P = 0.04$
12. Mental health evaluation	1 (3)	6 (15)	6 (13)	3.09; $P = 0.21$
13. Problem focused history	22 (65)	21 (53)	35 (73)	3.96; $P = 0.14$
14. Comprehensive history	22 (65)	30 (75)	39 (81)	2.88; $P = 0.24$
15. Problem focused exam	22 (65)	24 (60)	37 (77)	3.17; $P = 0.21$
16. Comprehensive physical	21 (62)	26 (65)	30 (63)	3.17; $P = 0.21$
17. Menstrual cycle info	14 (41)	21 (53)	30 (63)	4.11; $P = 0.13$
18. Pelvic exam	2 (6)	8 (25)	9 (19)	3.39; $P = 0.18$
19. Prescription of beta blockers	1 (3)	7 (18)	5 (10)	1.89; $P = 0.39$
20. Prescription of pain pills	10 (29)	19 (48)	25 (52)	4.40; $P = 0.11$
21. Reflex test	19 (56)	27 (68)	30 (63)	1.74; $P = 0.42$
22. Prophylactic therapy	7 (21)	12 (30)	15 (31)	1.26; $P = 0.53$
23. Hearing exam	8 (24)	12 (30)	18 (38)	1.85; $P = 0.40$
24. Visual screen	21 (62)	26 (65)	31 (65)	0.10; $P = 0.95$
25. Skin test	2 (6)	7 (18)	9 (19)	2.98; $P = 0.23$
26. Eye test	29 (85)	28 (70)	33 (69)	3.25; $P = 0.20$
27. MRI	11 (32)	11 (28)	11 (23)	0.90; $P = 0.64$
28. Beta strip	1 (3)	7 (18)	5 (10)	4.10; $P = 0.13$
29. Blood hormone levels	1 (3)	8 (25)	7 (15)	4.84; $P = 0.09$
30. CBC with diff	15 (44)	21 (53)	18 (38)	1.99; $P = 0.37$
31. Genetic counseling	1 (3)	4 (10)	4 (8)	1.45; $P = 0.49$
32. Pregnancy test	6 (18)	14 (35)	14 (65)	2.82; $P = 0.24$
33. Metabolic panel	14 (41)	21 (5)	25 (52)	1.21; $P = 0.55$
34. Blood typing	1 (3)	8 (25)	4 (8)	6.07; $P = 0.05$
35. X-ray	3 (9)	7 (18)	5 (10)	1.54; $P = 0.46$
36. Urinanalysis	9 (27)	13 (33)	13 (27)	0.43; $P = 0.81$
37. Ultrasound	0	5 (13)	4 (8)	4.31; $P = 0.12$
38. CT Scan	8 (24)	14 (35)	15 (31)	1.18; $P = 0.56$
39. Refer to neurologist	11 (32)	11 (28)	18 (38)	0.99; $P = 0.61$
40. Refer to cardiologist	1 (3)	4 (10)	4 (8)	1.45; $P = 0.49$
41. Preventative medicine consulation	4 (12)	5 (13)	6 (13)	0.01; $P = 0.99$

$F(1, 107) = 3.89$, $P = 0.05$; however, this difference was driven solely by the fact that three of the tests were relevant to women only (ie information about menstrual cycle, pregnancy test and pelvic exam) and the effect disappeared when these three items were removed [$F(1, 107) = 1.46$, $P = 0.23$]. Thus, gender of patient will not be discussed further.

A significant weight main effect on the Medical procedures form revealed that, in recommending procedures, physicians were influenced strongly by the weight of the patient, $F(1, 107) = 3.65$, $P < 0.03$. A linear trend analysis follow-up revealed the that most procedures were recommended for obese patients ($M = 12.18$), the second most for overweight patients ($M = 10.59$), and the least number for average weight patients ($M = 9.71$), $t(105) = 2.85$, $P < 0.01$.

Given that obese patients were presenting with two medical conditions, it is not necessarily surprising that more tests were recommended for obese than for average weight patients. In fact, it is a little surprising that a larger discrepancy was not found considering that many tests related medically or stereotypically to overweight and obesity. These included measures of cholesterol level, tryglycerides level, body fat percentages and glucose levels; dietary information and stress levels; prescribing anti-depressants; consulting about weight loss, an exercise program, and nutrition; and referring to a psychologist and obtaining a mental health evaluation. To specifically assess these items, a series of chi-square analyses with correction for multiple comparisons were conducted on these particular medical

procedures, as well as the non-weight-related medical procedures. As shown in Table 1, the results revealed significant differences on seven of the 12 weight-related procedures (see top half of Table 1), and only one of the 30 non-weight-related procedures (see bottom half of Table 1). All of the significant findings reveal the same pattern, namely that heavier patients were prescribed each of the tests more than were average-weight patients.

The results of the Patient follow-up questionnaire provided support for the notion that physicians viewed and responded to patients differently depending on their weight. The heavier the patients were, the more negative the attitudes and the distancing behaviors were. Such patterns can be observed from examining the means, the *F*s from ANOVAS, and the *t*s from the linear trend analyses, all of which are presented in Table 2. In particular, physicians reported that they would spend significantly less time with patients the heavier they were ($M = 31.13$ min with average-weight patients, $M = 25.00$ min with moderately overweight patients, and $M = 22.14$ min with severely overweight patients). On the additional 13 affective and behavioral items, 12 of the items were significant on both the overall *F*s and the linear trend analyses showing that, the heavier the patients were, the more negativity they faced.

Gender of physician

While the ratio of participants reflected an accurate representation of practicing female/male physicians in the Texas Medical Center, the distribution of female physicians into each of the six conditions was insufficient for adequately examining physician gender effects (eg, smallest cell size, $n = 1$). Further, the exploratory analyses that we attempted with these small *n*s revealed no significant gender effects.

Discussion

Responses to the two questionnaires reveal a great deal about physicians' perceptions of and intended courses of treatment for patients who are overweight and obese. The results from the medical procedures form reveal that physicians were just as likely to recommend non-weight-related tests and procedures to patients of varying weight, but were more likely to run weight-related tests on heavier patients. On one hand, physicians seem to provide appropriate levels of care to heavier patients by responding to both the headache and the weight. On the other hand, are physicians responding enough? If the patients are obese, more of a burden might be placed on the physician to speak about it, and to introduce and continue promoting discussions concerning weight loss,

Table 2 Effects of patient weight on physicians' attitudes and decisions of care

	Average, BMI = 23, n = 31 M; (s.d.)	*Overweight, BMI = 30, n = 36 M; (s.d.)*	*Obese, BMI = 36, n = 43 M; (s.d.)*	F*(2,107)*	P	t *(2,107)*	P
Time I would spend with patient (min)	31.1[a] (9.4)	25.0[b] (9.9)	22.4[c] (8.3)	8.38	0.00	15.88	0.00
1. How healthy is the patient	7.3[a] (1.1)	5.9[b] (1.8)	5.4[b] (1.8)	14.02	0.00	28.01	0.00
2. Patient takes care of himself/herself	6.8[a] (1.2)	5.4[b] (1.8)	4.5 (1.6)	20.97	0.00	41.94	0.00
3. Patient is self-disciplined	6.4[a] (1.5)	4.8[b] (1.8)	4.1[c] (1.5)	18.58	0.00	36.92	0.00
4. Level of stictness in the medical advice I'd give	5.0[a] (2.0)	6.4[b] (1.8)	6.0[c] (1.5)	6.44	0.00	7.20	0.01
5. Seriousness of the patient's health problem	4.6 (1.8)	5.1 (1.5)	4.7 (1.6)	1.20	0.31	0.22	0.64
6. Seeing this patient would feel like a waste of my time	2.0[a] (1.1)	2.3[a] (1.5)	3.2[b] (2.2)	4.58	0.01	8.13	0.01
7. This sort of patient would make me like my job	5.8[a] (2.7)	4.8 (2.5)	4.0[b] (2.0)	5.33	0.01	10.61	0.00
8. Amount of patience I would have	7.6[a] (1.2)	7.0 (1.2)	6.5[b] (1.5)	4.53	0.01	8.69	0.00
9. Extent to which this patient would annoy me	2.3[a] (1.8)	2.6[a] (1.9)	3.4[b] (1.9)	3.94	0.02	7.54	0.01
10. Personal desire I have to help this patient	7.7[a] (1.2)	7.2 (1.6)	6.7[b] (1.6)	3.67	0.03	7.30	0.01
11. Likelihood that the patient would follow my advice	7.2[a] (1.3)	6.3[b] (1.6)	5.4[c] (1.9)	11.06	0.00	22.03	0.00
12. I believe that patient would benefit from counseling	4.5[a] (2.1)	5.3 (2.1)	5.9[b] (2.2)	6.40	0.00	12.75	0.00
13. My overall positivity toward the patient	7.2[a] (1.4)	6.7 (1.7)	6.4[b] (1.6)	2.45	0.09	4.84	0.03

Note: all items were presented on nine-point scales that were anchored by (1) = 'Not at all', (5) = 'Somewhat', and (9) = 'Extremely'. Differences in superscripts across the rows reflect *post-hoc* comparisons of $P < 0.05$.

nutrition and exercise programs. In the current study, however, only 42% of physicians chose to discuss weight loss with obese patients (comparable to the 39% reported in a Risk Factor Surveillance System sample of 12 835 adults[23]) and 35% chose to discuss it with overweight patients. Similarly, only 31% of physicians indicated referring obese patients to a nutrition counselor and 30% indicated similarly referring overweight patients. Finally, only 27% of the physicians seeing obese patients (comparable to the 34% reported in a 1995 National Health Interview Survey with 17 317 respondents[24]) and 23% of those seeing overweight patients indicated that they would broach the topic of exercise. Past studies showing comparable percentages to those reported here suggest that physicians may need to take a greater role in aggressively addressing obesity (eg, increased physician counseling) in their patients.[23–25] Past studies suggest that such behaviors on the part of physicians might be promoted if they received respective reimbursement, did not have limited times during office visits, had more training in counseling, had greater confidence in their own ability to counsel, and had confidence that their patients could lose weight.[24]

While the results from the recommended medical procedures are not conclusive, the current results do suggest that physicians are responding to the stigma of obesity. For instance, the results show that physicians are more likely to recommend psychological counseling to heavier individuals, suggesting a belief that those who are overweight must also be unhappy and unstable. In line with this, 22 physicians included handwritten comments in the margins of the returned experimental packets suggesting such stigmatization. One physician wrote, in response to an obese patient, 'this woman has a very unhappy life'. Another indicated that the obese patient was 'most likely a drug addict'. A third indicated that the obese female patient was 'suffering underlying depression'. Additionally, on more than a dozen occasions, physicians suggested giving the overweight patients anti-depression medication.

Results from the Patient follow-up questionnaire reveal a very different, somewhat clearer picture of stigmatization against obesity. First, the heavier the patient was, the less time the physician proposed that he or she would spend with the patient. Considering that 31 min is the mean duration of a visit proposed for caring for average weight patients with a migraine headache, the respectively projected times of 25 min and 22 min for overweight and obese patients would be insufficient to care for the same migraine headache condition as well as the second condition involving weight. Less time may very likely result in less attention, less patient–physician interaction, and less individualized consideration of treatments. Certainly if physicians give additional tests (whether weight-related or not) to heavier patients, they may be giving compromised care — they are doing more tests in a much shorter period of time.

Other evidence from the Patient follow-up questionnaire reveals that physicians perceived heavier patients in congruence with the obesity stereotype.[6,7,10,26] As the patient got heavier, physicians judged them to be increasingly less *healthy*, worse in *taking care* of themselves, and less *self-disciplined*. Physicians also reported that they would need to give *stricter* medical advice to overweight and obese patients than average weight patients. The sum of these findings support the possibility that heavier patients actually are a greater burden to physicians because of their increased serious health risks. But, contrary to this ideology, it is interesting to note that when asked to judge the *seriousness* of the medical condition, no significant differences emerged as a function of weight. While certainly excess weight has been linked with increased health risks, the pattern of responses seems to reflect that physicians feel more negativity toward heavier patients. Specifically, Table 2 shows that physicians reported that seeing patients was a greater *waste of their time* the heavier that they were, that physicians would *like their jobs less* as their patients increased in size, that heavier patients were viewed to be more *annoying*, and that physicians felt less *patience* the heavier the patient was. It is important to note, however, that while these means do reflect statistically significant differences, physicians did not wildly disparage the heavier patients in absolute terms. That is, physicians felt that even the obese patients would not clearly be a waste of their time ($M = 2.03$ on a nine-point scale). However, there were still differences as a function of patients' weight. Perhaps the most disturbing pattern shown in Table 2 is that physicians indicated having significantly less *personal desire to help* patients the heavier they were.

In terms of outcomes, physicians predicted that heavier patients would be less likely to *follow their advice* and that heavier patients would *benefit from counseling*. While only marginally significant, they also reported feeling that the heavier the patient was, the less *positivity* overall they would feel toward the patient.

In sum, the quality of care that physicians reported that they would distribute to their patients was largely influenced by the weight of the patients. Importantly, the current results revealed a very strong and consistent linear trend in the way that physicians respond to the size of their patients. Namely, physicians perceived obese patients more negatively than they perceived overweight patients, and they perceived overweight patients more negatively than they perceived average-weight patients. Thus, this study provides the first known empirical demonstration indicating that physicians' perceptions worsen as the weight of a patient steadily increases.

Limitations and implications

An important question concerning these results is whether the attitudes and planned behaviors would translate into actual differences as the patient on the form becomes a real live patient that the physician encounters in the office. The comparisons that we were able to draw with already existent data shows large congruency in the extent to which

physicians intend to and actually talk to their obese patients about exercise[23] and weight.[24] Such consistent findings suggest some generalizability from the current results to actual behaviors. Similarly, past research suggests that attitudes and planned intentions are a particularly good indicator of actual behavior.[27,28] Thus, we believe our results are externally valid and that the climate and physician/patient connection might be lacking in warmth and positivity when patients are heavier. Physicians who are hostile or even mildly annoyed about overweight and obesity might be more likely to rush the medical appointment, be biased in the approach that they take and medical care that they give, or terminate the interaction before fully helping and serving the patient.

Physicians' negative attitudes also enact a self-fulfilling prophecy[29] whereby physicians treat overweight patients less favorably and overweight patients, as a result, take care of themselves less well. In addition, overweight patients may respond negatively to the physicians' negative attitudes, thereby reinforcing negative physician attitudes. If physicians' hopes for the patient are not favorable, if affectivity is not positive, and if the medical care is not up to par with what average weight individuals are receiving, overweight individuals might adopt lower standards for their personal care. Furthermore, it is possible that overweight people may delay or avoid altogether seeking the medical attention they need,[9,30] and this avoidance behavior may be a major contributing factor in elevating the mortality rate of overweight individuals.[31]

In conclusion, the results revealed considerable support for the hypothesis that physicians hold different beliefs about and propose that they would provide different levels of care (eg amount of time spent) to obese, overweight and average-weight individuals. In particular, physicians' attitudes and behaviors were significantly influenced by a patient's weight status such that the heavier individuals were, the more negatively they were perceived and treated. From a public health perspective, overweight patients are a vulnerable population not only because they are at higher health risk, but also because they are the targets of stereotypes and discrimination. The implications of this study suggest that the burden of improved healthcare for the overweight and obese patient rests not only on the responsibility of the patient but also on improving the attitudes and behaviors of their physicians.

Acknowledgements

The authors thank Dr Clifford Dasco and Dr David Lairson for comments on earlier copies of this manuscript. This study was supported by two mini-grants from the University of Texas-Houston School of Public Health and from funding by Rice University.

References

1 Flegal KM, Carroll MD, Kuczmarski RJ, Johnson CL. Overweight and overweight in the United States: prevalence and trends 1960–1994. *Int J Obes Relat Metab Disord* 1998; **22**: 39–47.
2 Mokdad AH, Serdula MK, Dietz WH, Bowman BA, Marks JS, Koplan JP. The spread of the obesity epidemic in the United States 1991–1998. *JAMA* 1999; **282**: 1519–1522.
3 Must A, Strauss RS. Risk and consequences of childhood and adolescent overweight. *Int J Obes Relat Metab Disord* 1999; **23**: S2–S11.
4 Pi-Sunyer FX. Health implications of overweight. *Am J Clin Nutr* 1991; **53**: 1595S–1603S.
5 Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB. Annual deaths attributable to obesity in the United States. *JAMA* 1999; **282**: 1530–1538.
6 DeJong W, Kleck RE. The social psychological effects of overweight. In Herman, CP, Zanna MP, Higgins ET (eds). *Physical appearance stigma and social behavior: the Ontario Symposium*. Erlbaum: New Jersey; 1986. pp 65–87.
7 Hill AJ, Williams J. Psychological health in a non-clinical sample of obese women. *Int J Obes Relat Metab Disord* 1998; **22**: 578–583.
8 Rothblum ED. The stigma of women's weight: social and economic realities. *Femin Psychol* 1992; **21**: 61–73.
9 Allon N. The stigma of overweight in everyday life. In: Wolman BB, DeBerry S (eds). *Psychological aspects of overweight: a handbook*. Van Nostrand Reinhold: New York; 1982. pp 130–174.
10 Weiner B. *Judgements of responsibility: a theory of social conduct*. The Guilford Press: New York; 1995.
11 Pingitore R, Dugoni BL, Tindale RS, Spring B. Bias against overweight job applicants in a simulated employment interview. *J Appl Psychol* 1994; **79**: 909–917.
12 Roehling MV. Weight-based discrimination in employment: psychological and legal aspects. *Pers Psychol* 1999; **52**: 969–1016.
13 Canning H, Mayer J. Obesity: an influence on high school performance? *Am J Clin Nutr* 1967; **20**: 352–354.
14 Hebl MR, Heatherton TF. The stigma of obesity: the differences are black and white. *Pers Soc Psychol Bull* 1997; **24**: 417–426.
15 Pérusse L, Chagnon YC, Weisnagel J, Bouchard C. The human obesity gene map: the 1998 update. *Obes Res* 1999; **71**: 111–129.
16 Stunkard AJ, Harris JR, Pedersen NL, McClearn GE. The body-mass index of twins who have been reared apart. *New Engl J Med* 1990; **322**: 1483–1487.
17 Wiese HJC, Wilson JF, Jones RA, Neises, M. Obesity stigma reduction in medical students. *Int J Obes Relat Metab Disord* 1992; **16**: 859–868.
18 Young LM, Powell B. The effects of overweight on the clinical judgements of mental health professionals. *J Health Soc Behav* 1985; **26**: 233–246.
19 Price JH, Desmond SM, Krol RA, Snyder FF, O'Connell JK. Family practice physicians' beliefs attitudes and practices regarding overweight. *Am J Prev Med* 1987; **36**: 339–345.
20 Bretytspraak LM, McGee J, Conger JC, Whatley JL, Moore JT. Sensitizing medical students to impression formation processes in the patient interview. *J Med Educ* 1977; 47–54.
21 Blumberg P, Mellis LP. Medical students' attitudes toward the obese and the morbidly obese. *Int J Eat Disord* 1985; **4**: 169–175.
22 Maddox G, Liederman V. Overweight as a social disability with medical implications. *J Med Educ* 1969; **44**: 214–220.
23 Wee CC, McCarthy EP, Davis RB, Phillips RS. Physician counseling about exercise. *JAMA* 1999; **282**: 1583–1588.
24 Galuska DA, Will JC, Serdula MK, Ford ES. Are health care professionals advising obese patients to lose weight? *JAMA* 1999; **282**: 1576–1578.
25 Fontanarosa PB. Patients, physicians, and weight control. *JAMA* 1999; **282**: 1581–1582.
26 Harris MB, Harris RJ, Bochner S. Fat four-eyed and female: stereotypes of obesity, glasses and gender. *J Appl Soc Psychol* 1982; **126**: 503–516.

27 Ajzen I, Fishbein M. Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychol Bull* 1977; **84**: 888–918.
28 Eagly A, Chaiken S. *The psychology of attitudes*. HBJ; 1993.
29 Snyder M, Tanke ED, Berscheid E. Social perception and interpersonal behavior: on the self-fulfilling nature of social stereotypes. *J Pers Soc Psychol* 1977; **35**: 656–666.
30 Adams CH, Smith NJ, Wilbur DC, Grady KE. The relationship of overweight to the frequency of pelvic examinations: do physician and patient attitudes make a difference? *Women Health* 1993; **202**: 45–57.
31 Abe R, Kumagai MK, Hirosaki A, Nakamura T. Biological characteristics of breast cancer in overweight. *Tohoku J Exp Med* 1976; **120**: 351–359.

NEXTIONS

THINK SMARTER | LEAD BETTER

Yellow Paper Series

Written in Black & White

Exploring Confirmation Bias in Racialized Perceptions of Writing Skills

Lead Researcher

Dr. Arin N. Reeves

2014-0404

RESEARCH QUESTION: *Given our finding in a previous study that supervising lawyers are more likely than not to perceive African American lawyers as having subpar writing skills in comparison to their Caucasian counterparts, we asked if confirmation bias unconsciously causes supervising lawyers to more negatively evaluate legal writing by an African American lawyer.*

CONFIRMATION BIAS:

A mental shortcut – a bias – engaged by the brain that makes one actively seek information, interpretation and memory to only observe and absorb that which affirms established beliefs while missing data that contradicts established beliefs.

We first discovered empirical evidence that supervising lawyers perceived African Americans lawyers to be subpar in their writing skills in comparison to their Caucasian counterparts when we researched unconscious biases in the legal profession over ten years ago. Since our surveys and focus groups at the time were studying unconscious biases generally, we decided to study this specific bias of writing skills in greater detail via the cognitive construct of **confirmation bias.**

This research summary provides a general overview of the methodology, results and key takeaways from the study. Please note that we studied this question only from the unconscious or implicit bias perspective. While the possibility of explicit bias exists, our research has consistently shown that implicit bias is far more prevalent in our workplaces today than explicit bias, thereby guiding us to utilize our resources to study implicit instead of explicit biases.

Methodology

Nextions, along with the assistance of 5 partners from 5 different law firms, drafted a research memo from a hypothetical third year litigation associate that focused on the issue of trade secrets in internet start-ups. We followed a simple Question Presented, Brief Answer, Facts, Discussion and Conclusion format for the memo, and we deliberately inserted 22 different errors, 7 of which were minor spelling/grammar errors, 6 of which were substantive technical writing errors, 5 of which were errors in fact, and 4 of which were errors in the analysis of the facts in the Discussion and Conclusion sections.

This memo was then distributed to 60 different partners (who had previously agreed to participate in a "writing analysis study" from 22 different law firms of whom 23 were women, 37 were men, 21 were racial/ethnic minorities, and 39 were Caucasian. While all of the partners received the same memo, half the partners received a memo that stated the associate was African American while the other half received a memo that stated the associate was Caucasian:

While all of the partners received the same memo, half the partners received a memo that stated the associate was African American while the other half received a memo that stated the associate was Caucasian.

Name: *Thomas Meyer*	***Name:*** *Thomas Meyer*
Seniority: *3rd Year Associate*	***Seniority:*** *3rd Year Associate*
Alma Mater: *NYU Law School*	***Alma Mater:*** *NYU Law School*
Race/Ethnicity: *African American*	***Race/Ethnicity:*** *Caucasian*

The 60 partners in the study received the memo electronically (an attached pdf) along with the research materials used in the preparation of the memo. The cover email thanked each of them for participating in a study on "writing competencies of young attorneys," and asked them to edit the memo for all factual, technical and substantive errors. The partners were also asked to rate the overall quality of the memo from a 1 to 5, with "1" indicating the memo was extremely poorly written and "5" extremely well written.

The partners were originally given 4 weeks to complete the editing and rating, but we had to extend deadline to 7 weeks in order to obtain more responses. 53 partners completed the editing and rating of the memo. Of the 53 completed responses, 24 had received the memo by the "African American" Thomas Meyer, and 29 had received the memo by the "Caucasian" Thomas.

General Findings

The exact same memo, averaged a 3.2/5.0 rating under our hypothetical "African American" Thomas Meyer and a 4.1/5.0 rating under hypothetical "Caucasian" Thomas Meyer.

The exact same memo, averaged a 3.2/5.0 rating under our hypothetical "African American" Thomas Meyer and a 4.1/5.0 rating under hypothetical "Caucasian" Thomas Meyer. The qualitative comments on memos, consistently, were also more positive for the "Caucasian" Thomas Meyer than our "African American" Thomas Meyer:

"Caucasian" Thomas Meyer	**"African American" Thomas Meyer**
"generally good writer but needs to work on…"	*"needs lots of work"*
	"can't believe he went to NYU"
"has potential"	
"good analytical skills"	*"average at best"*

In regards to the specific errors in the memo:

- An average of 2.9/7.0 spelling grammar errors were found in "Caucasian" Thomas Meyer's memo in comparison to 5.8/7.0 spelling/grammar errors found in "African American" Thomas Meyer's memo.
- An average of 4.1/6.0 technical writing errors were found in "Caucasian" Thomas Meyer's memo in comparison to 4.9/6.0 technical writing errors found in "African American" Thomas Meyer's memo.
- An average of 3.2/5.0 errors in facts were found in "Caucasian" Thomas Meyer's memo in comparison to 3.9/5.0 errors in facts were found in "African American" Thomas Meyer's memo.

The 4 errors in analysis were difficult to parse out quantitatively because of the variances in narrative provided by the partners as to why they were analyzing the writing to contain analytical errors. Overall though, "Caucasian" Thomas Meyer's memo was evaluated to be better in regards to the analysis of facts and had substantively fewer critical comments.

General Findings Cont.

We did not ask for edits and/or comments on formatting. However, we did receive such edits and/or comments in 41 out of the 53 responses, and all of them regarded changes that the partners would have liked to see on the formatting in the memo. Of the 41 edits and/or comments on formatting, 11 were for "Caucasian" Thomas Meyer's memo in comparison to 29 for "African American" Thomas Meyer's memo.

There was no significant correlation between a partner's race/ethnicity and the differentiated patterns of errors found between the two memos. There was also no significant correlation between a partner's gender and the differentiated patterns of errors found between the two memos. We did find that female partners generally found more errors and wrote longer narratives than the male partners.

Analysis & Discussion

We undertook this study with the hypothesis that unconscious confirmation bias in a supervising lawyer's assessment of legal writing would result in a more negative rating if that writing was submitted by an African American lawyer in comparison to the same submission by a Caucasian lawyer. In order to create a study where we could control for enough variables to truly see the impact of confirmation bias, we did not study the potential variances that can be caused due to the intersection of race/ethnicity, gender, generational differences and other such salient identities. Thus, our conclusion is limited to the impact of confirmation bias in the evaluation of African American men in comparison to Caucasian men. We do not know (although we plan to study the issue in the very near future!) how this impact will splinter or strengthen when gender and/or other identities are introduced.

Confirmation bias manifests itself most often in the "data gathering" phase of our evaluation – the time during which we seek out errors, and this manifestation is almost always unconscious.

The data findings affirmed our hypothesis, but they also illustrated that the confirmation bias on the part of the evaluators occurred in the data collection phase of their evaluation processes – the identification of the errors – and not the final analysis phase. When expecting to find fewer errors, we find fewer errors. When expecting to find more errors, we find more errors. That is unconscious confirmation bias. Our evaluators unconsciously found more of the errors in the "African American" Thomas Meyer's memo, but the final rating process was a conscious and unbiased analysis based on the number of errors found. When partners say that they are evaluating assignments without bias, they are probably right in believing that there is no bias in the assessment of the errors found; however, if there is bias in the finding of the errors, even a fair final analysis cannot, and will not, result in a fair result.

Key Takeaways

There are commonly held racially-based perceptions about writing ability that unconsciously impact our ability to objectively evaluate a lawyer's writing… These commonly held perceptions translate into confirmation bias in ways that impact what we see as we evaluate legal writing. We see more errors when we expect to see errors, and we see fewer errors when we do not expect to see errors.

There are commonly held racially-based perceptions about writing ability that unconsciously impact our ability to objectively evaluate a lawyer's writing. Most of the perceptions uncovered in research thus far indicate that commonly held perceptions are biased against African Americans and in favor of Caucasians.

These commonly held perceptions translate into confirmation bias in ways that impact what we see as we evaluate legal writing. We see more errors when we expect to see errors, and we see fewer errors when we do not expect to see errors.

Recommendations for Next Actions

Infusing the point at which unconscious thought has greatest impact with objective mechanisms that force the conscious brain to add input, decreases unconscious bias greatly. We have worked with many employers to revise their formal and informal evaluation processes to be more infused with objective interrupters that compel unconscious biases to be filtered through conscious analysis, and we have seen many success stories. **So, make the subjective more objective in order to make the unconscious more conscious.**

EXAMPLE: In one law firm where we found that minority summer associates were consistently being evaluated more negatively than their majority counterparts, we created an interruption mechanism to infuse the subjective with objective. We worked with the firm to create an Assignment Committee, comprised of 3 partners through whom certain assignments were distributed to the summer associates and through whom the summer associates submitted work back to the partners who needed the work done. When the work was evaluated, the partners evaluating the work did not know which associate had completed the work. The assignments for this process were chosen judiciously, and there was a lot of work done to ensure buy-in from all partners. At the end of the summer, every associate had at least 2 assignments that had been graded blindly. The firm then examined how the blind evaluations compared with the rest of the associate's evaluations and found that the blind evaluations were generally more positive for minorities and women and less positive for majority men.

Ideas for Inclusion

- Distribute and discuss this study with senior lawyers in your organization to gather their reactions and perspectives. Ask them how they would recommend making the subjective more objective in order to reduce confirmation bias in their evaluation processes.

- If racial/ethnic minorities are deemed to be subpar in writing skills, send out samples of a minority lawyer's writing and a sample of a majority lawyer's writing without any identifying information attached. Ask a few senior lawyers to evaluate both samples. Explore how the samples may be evaluated differently when the lawyer's background is not available.

- Implement training on unconscious bias for everyone who is in an evaluative position. Our unconscious bias trainings have proven effective in reducing bias through raising awareness and insights into how unconscious biases operate and can be interrupted.

- If you offer writing assistance in the form of coaches, workshops and such, offer the assistance to everyone, not just racial/ethnic minorities in order to prevent the reification of the bias.

Distribute and discuss this study with senior lawyers in your organization to gather their reactions and perspectives.

Lead Researcher:

Dr. Arin N. Reeves | 312.922.0226

NEXTIONS

THINK SMARTER | LEAD BETTER

500 N. Michigan Ave., Suite 600
Chicago, Illinois 60611
P 312 922.0226

Race & Ethnic Fairness in the Courts

Implicit Bias

A Primer for Courts

Jerry Kang

Prepared for the National Campaign to Ensure the Racial and Ethnic Fairness of America's State Courts

August 2009

ABOUT THE PRIMER

This Primer was produced as part of the National Campaign to Ensure the Racial and Ethnic Fairness of America's State Courts. The Campaign seeks to mobilize the significant expertise, experience, and commitment of state court judges and court officers to ensure both the perception and reality of racial and ethnic fairness across the nation's state courts. The Campaign is funded by the Open Society Institute, the State Justice Institute, and the National Center for State Courts. Points of view or opinions expressed in the Primer are those of the author and do not represent the official position of the funding agencies. To learn more about the Campaign, visit www.ncsconline.org/ref.

ABOUT THE AUTHOR & REVIEWERS

Jerry Kang is Professor of Law at UCLA School of Law. He has written and lectured extensively on the role of implicit bias in the law. For more information on Professor Kang, please visit jerrykang.net. The Primer benefited from the review and comments of several individuals working with the National Campaign, including Dr. Pamela Casey, Dr. Fred Cheesman, Hon. Ken M. Kawaichi, Hon. Robert Lowenbach, Dr. Shawn Marsh, Hon. Patricia M. Martin, Ms. Kimberly Papillon, Hon. Louis Trosch, and Hon. Roger K. Warren.

Table of Contents

Implicit Bias: A Primer

Schemas and Implicit Cognitions (or "mental shortcuts")

Stop for a moment and consider what bombards your senses every day. Think about everything you see, both still and moving, with all their color, detail, and depth. Think about what you hear in the background, perhaps a song on the radio, as you decode lyrics and musical notes. Think about touch, smell, and even taste. And while all that's happening, you might be walking or driving down the street, avoiding pedestrians and cars, chewing gum, digesting your breakfast, flipping through email on your smartphone. How does your brain do all this simultaneously?

It does so by processing through schemas, which are templates of knowledge that help us organize specific examples into broader categories. When we see, for example, something with a flat seat, a back, and some legs, we recognize it as a "chair." Regardless of whether it is plush or wooden, with wheels or bolted down, we know what to do with an object that fits into the category "chair." Without spending a lot of mental energy, we simply sit. Of course, if for some reason we have to study the chair carefully--because we like the style or think it might collapse--we can and will do so. But typically, we just sit down.

We have schemas not only for objects, but also processes, such as how to order food at a restaurant. Without much explanation, we know what it means when a smiling person hands us laminated paper with detailed descriptions of food and prices. Even when we land in a foreign airport, we know how to follow the crazy mess of arrows and baggage icons toward ground transportation.

These schemas are helpful because they allow us to operate without expending valuable mental resources. In fact, unless something goes wrong, these thoughts take place automatically without our awareness or conscious direction. In this way, most cognitions are implicit.

Implicit Social Cognitions (or "thoughts about people you didn't know you had")

What is interesting is that schemas apply not only to objects (e.g., "chairs") or behaviors (e.g., "ordering food") but also to human beings (e.g., "the elderly"). We naturally assign people into various social categories divided by salient and chronically accessible traits, such as age, gender, race, and role. And just as we might have implicit cognitions that help us walk and drive, we have implicit social cognitions that guide our thinking about social categories. Where do these schemas come from? They come from our experiences with other people, some of them direct (i.e., real-world encounters) but most of them vicarious (i.e., relayed to us through stories, books, movies, media, and culture).

If we unpack these schemas further, we see that some of the underlying cognitions include stereotypes, which are simply traits that we associate with a category. For instance, if we think that a particular category of human beings is frail--such as the elderly--we will not raise our guard. If we think that another category is foreign--such as Asians--we will be surprised by their fluent English. These cognitions also include attitudes, which are overall, evaluative feelings that are positive or negative. For instance, if we identify someone as having graduated from our beloved alma mater, we will feel more at ease. The term "implicit bias"

includes both implicit stereotypes and implicit attitudes.

Though our shorthand schemas of people may be helpful in some situations, they also can lead to discriminatory behaviors if we are not careful. Given the critical importance of exercising fairness and equality in the court system, lawyers, judges, jurors, and staff should be particularly concerned about identifying such possibilities. Do we, for instance, associate aggressiveness with Black men, such that we see them as more likely to have started the fight than to have responded in self-defense? Or have we already internalized the lessons of Martin Luther King, Jr. and navigate life in a perfectly "colorblind" (or gender-blind, ethnicity-blind, class-blind, etc.) way?

Asking about Bias (or "it's murky in here")

One way to find out about implicit bias is simply to ask people. However, in a post-civil rights environment, it has become much less useful to ask explicit questions on sensitive topics. We run into a "willing and able" problem.

First, people may not be willing to tell pollsters and researchers what they really feel. They may be chilled by an air of political correctness.

Second, and more important, people may not know what is inside their heads. Indeed, a wealth of cognitive psychology has demonstrated that we are lousy at introspection. For example, slight environmental changes alter our judgments and behavior without our realizing. If the room smells of Lysol, people eat more neatly. People holding a warm cup of coffee (versus a cold cup) ascribe warmer (versus cooler) personality traits to a stranger described in a vignette. The experiments go on and on. And recall that by definition, implicit biases are those that we carry without awareness or conscious direction. So how do we know whether we are being biased or fair-and-square?

Implicit measurement devices (or "don't tell me how much you weigh, just get on the scale")

In response, social and cognitive psychologists with neuroscientists have tried to develop instruments that measure stereotypes and attitudes, without having to rely on potentially untrustworthy self-reports. Some instruments have been linguistic, asking folks to write out sentences to describe a certain scene from a newspaper article. It turns out that if someone engages in stereotypical behavior, we just describe what happened. If it is counter-typical, we feel a need to explain what happened. (Von Hippel 1997; Sekaquaptewa 2003).

Others are physiological, measuring how much we sweat, how our blood pressure changes, or even which regions of our brain light up on an fMRI (functional magnetic resonance imaging) scan. (Phelps 2000).

Still other techniques borrow from marketers. For instance, conjoint analysis asks people to give an overall evaluation to slightly different product bundles (e.g., how do you compare a 17" screen laptop with 2GB memory and 3 USB ports, versus a 15" laptop with 3 GB of memory and 2 USB ports). By offering multiple rounds of choices, one can get a measure of how important each feature is to a person even if she had no clue to the question "How much would you pay for an extra USB port?" Recently, social cognitionists have adapted this methodology by creating "bundles" that include demographic attributes. For instance, how

would you rank a job with the title Assistant Manager that paid $160,000 in Miami working for Ms. Smith, as compared to another job with the title Vice President that paid $150,000 in Chicago for Mr. Jones? (Caruso 2009).

Scientists have been endlessly creative, but so far, the most widely accepted instruments have used reaction times--some variant of which has been used for over a century to study psychological phenomena. These instruments draw on the basic insight that any two concepts that are closely associated in our minds should be easier to sort together. If you hear the word "moon," and I then ask you to think of a laundry detergent, then "Tide" might come more quickly to mind. If the word "RED" is painted in the color red, we will be faster in stating its color than the case when the word "GREEN" is painted in red.

Although there are various reaction time measures, the most thoroughly tested one is the Implicit Association Test (IAT). It is a sort of video game you play, typically on a computer, where you are asked to sort categories of pictures and words. For example, in the Black-White race attitude test, you sort pictures of European American faces and African American faces, Good words and Bad words in front of a computer. It turns out that most of us respond more quickly when the European American face and Good words are assigned to the same key (and African American face and Bad words are assigned to the other key), as compared to when the European American face and Bad words are assigned to the same key (and African American face and Good words are assigned to the other key). This average time differential is the measure of implicit bias. [If the description is hard to follow, try an IAT yourself at Project Implicit.]

Pervasive implicit bias (or "it ain't no accident")

It may seem silly to measure bias by playing a sorting game (i.e. the IAT). But, a decade of research using the IAT reveals pervasive reaction time differences in every country tested, in the direction consistent with the general social hierarchies: German over Turk (in Germany), Japanese over Korean (for Japanese), White over Black, men over women (on the stereotype of "career" versus "family"), light-skinned over dark skin, youth over elderly, straight over gay, etc. These time differentials, which are taken to be a measure of implicit bias, are systematic and pervasive. They are statistically significant and not due to random chance variations in measurements.

These pervasive results do not mean that everyone has the exact same bias scores. Instead, there is wide variability among individuals. Further, the social category you belong to can influence what sorts of biases you are likely to have. For example, although most Whites (and Asians, Latinos, and American Indians) show an implicit attitude in favor of Whites over Blacks, African Americans show no such preference on average. (This means, of course, that about half of African Americans do prefer Whites, but the other half prefer Blacks.)

Interestingly, implicit biases are dissociated from explicit biases. In other words, they are related to but differ sometimes substantially from explicit biases--those stereotypes and attitudes that we expressly self-report on surveys. The best understanding is that implicit and explicit biases are related but different mental constructs. Neither kind should be viewed as the solely "accurate" or "authentic" measure of bias. Both measures tell us something important.

Real-world consequences (or "why should we care?")

All these scientific measures are intellectually interesting, but lawyers care most about real-world consequences. Do these measures of implicit bias predict an individual's behaviors or decisions? Do milliseconds really matter>? (Chugh 2004). If, for example, well-intentioned people committed to being "fair and square" are not influenced by these implicit biases, then who cares about silly video game results?

There is increasing evidence that implicit biases, as measured by the IAT, do predict behavior in the real world--in ways that can have real effects on real lives. Prof. John Jost (NYU, psychology) and colleagues have provided a recent literature review (in press) of ten studies that managers should not ignore. Among the findings from various laboratories are:

- implicit bias predicts the rate of callback interviews (Rooth 2007, based on implicit stereotype in Sweden that Arabs are lazy);
- implicit bias predicts awkward body language (McConnell & Leibold 2001), which could influence whether folks feel that they are being treated fairly or courteously;
- implicit bias predicts how we read the friendliness of facial expressions (Hugenberg & Bodenhausen 2003);
- implicit bias predicts more negative evaluations of ambiguous actions by an African American (Rudman & Lee 2002), which could influence decisionmaking in hard cases;
- implicit bias predicts more negative evaluations of agentic (i.e. confident, aggressive, ambitious) women in certain hiring conditions (Rudman & Glick 2001);
- implicit bias predicts the amount of shooter bias--how much easier it is to shoot African Americans compared to Whites in a videogame simulation (Glaser & Knowles 2008);
- implicit bias predicts voting behavior in Italy (Arcari 2008);
- implicit bias predicts binge-drinking (Ostafin & Palfai 2006), suicide ideation (Nock & Banaji 2007), and sexual attraction to children (Gray 2005).

With any new scientific field, there remain questions and criticisms--sometimes strident. (Arkes & Tetlock 2004; Mitchell & Tetlock 2006). And on-the-merits skepticism should be encouraged as the hallmark of good, rigorous science. But most scientists studying implicit bias find the accumulating evidence persuasive. For instance, a recent meta-analysis of 122 research reports, involving a total of14,900 subjects, revealed that in the sensitive domains of stereotyping and prejudice, implicit bias IAT scores better predict behavior than explicit self-reports. (Greenwald et al. 2009).

And again, even though much of the recent research focus is on the IAT, other instruments and experimental methods have corroborated the existence of implicit biases with real world consequences. For example, a few studies have demonstrated that criminal defendants with more Afro-centric facial features receive in certain contexts more severe criminal punishment (Banks et al. 2006; Blair 2004).

Malleability (or "is there any good news?")

The findings of real-world consequence are disturbing for all of us who sincerely believe that we do not let biases prevalent in our culture infect our individual decisionmaking. Even a little bit. Fortunately, there is evidence

that implicit biases are malleable and can be changed.

- An individual's motivation to be fair does matter. But we must first believe that there's a potential problem before we try to fix it.
- The environment seems to matter. Social contact across social groups seems to have a positive effect not only on explicit attitudes but also implicit ones.
- Third, environmental exposure to countertypical exemplars who function as "debiasing agents" seems to decrease our bias.
    - In one study, a mental imagery exercise of imagining a professional business woman (versus a Caribbean vacation) decreased implicit stereotypes of women. (Blair et al. 2001).
    - Exposure to "positive" exemplars, such as Tiger Woods and Martin Luther King in a history questionnaire, decreased implicit bias against Blacks. (Dasgupta & Greenwald 2001).
    - Contact with female professors and deans decreased implicit bias against women for college-aged women. (Dasgupta & Asgari 2004).
- Fourth, various procedural changes can disrupt the link between implicit bias and discriminatory behavior.
    - In a simple example, orchestras started using a blind screen in auditioning new musicians; afterwards women had much greater success. (Goldin & Rouse 2000).
    - In another example, by committing beforehand to merit criteria (is book smarts or street smarts more important?), there was less gender discrimination in hiring a police chief. (Uhlmann & Cohen 2005).
    - In order to check against bias in any particular situation, we must often recognize that race, gender, sexual orientation, and other social categories may be influencing decisionmaking. This recognition is the opposite of various forms of "blindness" (e.g., color-blindness).

In outlining these findings of malleability, we do not mean to be Pollyanish. For example, mere social contact is not a panacea since psychologists have emphasized that certain conditions are important to decreasing prejudice (e.g., interaction on equal terms; repeated, non-trivial cooperation). Also, fleeting exposure to countertypical exemplars may be drowned out by repeated exposure to more typical stereotypes from the media (Kang 2005).

Even if we are skeptical, the bottom line is that there's no justification for throwing our hands up in resignation. Certainly the science doesn't require us to. Although the task is challenging, we can make real improvements in our goal toward justice and fairness.

The big picture (or "what it means to be a faithful steward of the judicial system")

It's important to keep an eye on the big picture. The focus on implicit bias does not address the existence and impact of explicit bias--the stereotypes and attitudes that folks recognize and embrace. Also, the past has an inertia that has not dissipated. Even if all explicit and implicit biases were wiped away through some magical wand, life today would still bear the burdens of an unjust yesterday. That said, as careful stewards of the justice system, we

should still strive to take all forms of bias seriously, including implicit bias.

After all, Americans view the court system as the single institution that is most unbiased, impartial, fair, and just. Yet, a typical trial courtroom setting mixes together many people, often strangers, from different social backgrounds, in intense, stressful, emotional, and sometimes hostile contexts. In such environments, a complex jumble of implicit and explicit biases will inevitably be at play. It is the primary responsibility of the judge and other court staff to manage this complex and bias-rich social situation to the end that fairness and justice be done--and be seen to be done.

Glossary

Note: Many of these definitions draw from Jerry Kang & Kristin Lane, A Future History of Law and Implicit Social Cognition (unpublished manuscript 2009)

Attitude

An attitude is "an association between a given object and a given evaluative category." R.H. Fazio, et al., Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association, 18 J. EXPERIMENTAL SOCIAL PSYCHOLOGY 339, 341 (1982). Evaluative categories are either positive or negative, and as such, attitudes reflect what we like and dislike, favor and disfavor, approach and avoid. See also stereotype.

Behavioral realism

A school of thought within legal scholarship that calls for more accurate and realistic models of human decision-making and behavior to be incorporated into law and policy. It involves a three step process:

First, identify advances in the mind and behavioral sciences that provide a more accurate model of human cognition and behavior.

Second, compare that new model with the latent theories of human behavior and decision-making embedded within the law. These latent theories typically reflect "common sense" based on naïve psychological theories.

Third, when the new model and the latent theories are discrepant, ask lawmakers and legal institutions to account for this disparity. An accounting requires either altering the law to comport with more accurate models of thinking and behavior or providing a transparent explanation of "the prudential, economic, political, or religious reasons for retaining a less accurate and outdated view." Kristin Lane, Jerry Kang, & Mahzarin Banaji, Implicit Social Cognition and the Law, 3 ANNU. REV. LAW SOC. SCI. 19.1-19.25 (2007)

Dissociation

Dissociation is the gap between explicit and implicit biases. Typically, implicit biases are larger, as measured in standardized units, than explicit biases. Often, our explicit biases may be close to zero even though our implicit biases are larger.

There seems to be some moderate-strength relation between explicit and implicit biases. See Wilhelm Hofmann, A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures, 31 PERSONALITY & SOC. PSYCH. BULL. 1369 (2005) (reporting mean population correlation r=0.24 after analyzing 126 correlations). Most scientists reject the idea that implicit biases are the only "true" or "authentic" measure; both explicit and implicit biases contribute to a full understanding of bias.

Explicit

Explicit means that we are aware that we have a particular thought or feeling. The term sometimes also connotes that we have an accurate understanding of the source of that thought or feeling. Finally, the term often connotes conscious endorsement of the thought or feeling. For example, if one has an explicitly positive attitude toward chocolate, then one has a positive attitude, knows that one has a positive attitude, and consciously endorses and celebrates that preference. See also implicit.

Implicit

Implicit means that we are either unaware of or mistaken about the source of the thought or feeling. R. Zajonc, Feeling and thinking: Preferences need no inferences, 35 AMERICAN PSYCHOLOGIST 151 (1980). If we are unaware of a thought or feeling, then we cannot report it when asked. See also explicit.

Implicit Association Test

The IAT requires participants to classify rapidly individual stimuli into one of four distinct categories using only two responses (for example, in a the traditional computerized IAT, participants might respond using only the "E" key on the left side of the keyboard, or "I" on the right side). For instance, in an age attitude IAT, there are two social categories, YOUNG and OLD, and two attitudinal categories, GOOD and BAD. YOUNG and OLD might be represented by black-and-white photographs of the faces of young and old people. GOOD and BAD could be represented by words that are easily identified as being linked to positive or negative affect, such as "joy" or "agony". A person with a negative implicit attitude toward OLD would be expected to go more quickly when OLD and BAD share one key, and YOUNG and GOOD the other, than when the pairings of good and bad are switched.

The IAT was invented by Anthony Greenwald and colleagues in the mid 1990s. Project Implicit, which allows individuals to take these tests online, is maintained by Anthony Greenwald (Washington), Mahzarin Banaji (Harvard), and Brian Nosek (Virginia).

Implicit Attitudes

"Implicit attitudes are introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects." Anthony Greenwald & Mahzarin Banaji, Implicit social cognition: attitudes, self-esteem, and stereotypes, 102 Psychol. Rev. 4, 8 (1995). Generally, we are unaware of our implicit attitudes and may not endorse them upon self-reflection. See also attitude; implicit.

Implicit Biases

A bias is a departure from some point that has been marked as "neutral." Biases in implicit stereotypes and implicit attitudes are called "implicit biases."

Implicit Stereotypes

"Implicit stereotypes are the introspectively unidentified (or inaccurately identified) traces of past experience that mediate attributions of qualities to members of a social category" Anthony Greenwald & Mahzarin Banaji, Implicit social cognition: attitudes, self-esteem, and stereotypes, 102 Psychol. Rev. 4, 8 (1995). Generally, we are unaware of our implicit stereotypes and may not endorse them upon self-reflection. See also stereotype; implicit.

Implicit Social Cognitions

Social cognitions are stereotypes and attitudes about social categories (e.g., Whites, youths, women). Implicit social cognitions are implicit stereotypes and implicit attitudes about social categories.

Stereotype

A stereotype is an association between a given object and a specific attribute. An example is "Norwegians are tall." Stereotypes may support an overall attitude. For instance, if one likes tall people and Norwegians are tall, it is likely that this attribute will contribute toward a positive orientation toward Norwegians. See also attitude.

Validities

To decide whether some new instrument and findings are valid, scientists often look for various validities, such as statistical conclusion validity, internal validity, construct validity, and predictive validity.

- Statistical conclusion validity asks whether the correlation is found between independent and dependent variables have been correctly computed.
- Internal validity examines whether in addition to correlation, there has been a demonstration of causation. In particular, could there be potential confounds that produced the correlation?
- Construct validity examines whether the concrete observables (the scores registered by some instrument) actually represent the abstract mental construct that we are interested in. As applied to the IAT, one could ask whether the test actually measures the strength of mental associations held by an individual between the social category and an attitude or stereotype
- Predictive validity examines whether some test predicts behavior, for example, in the form of evaluation, judgment, physical movement or response. If predictive validity is demonstrated in realistic settings, there is greater reason to take the measures seriously.

Bibliography

Allen R. McConnell & Jill M. Leibold, Relations Among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes, 37 J. Experimental Soc. Psychol. 435 (2001)

Anthony G. Greenwald, et al., Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity, 97 J. Personality & Soc. Psychol. 17 (2009)

Brian D. Ostafin, & Tibor P. Palfai, Compelled to Consume: The Implicit Association Test and Automatic Alcohol Motivation, 20 Psych. of Addictive Behav. 322 (2006)

Claudia Goldin & Cecilia Rouse, Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians, 90 Am. Econ. Rev. 715 (2000)

Dan-Olof Rooth, Implicit Discrimination in Hiring: Real World Evidence (Institute for the Study of Labor, Discussion Paper No. 2764, April 2007)

Denise Sekaquaptewa, et al., Stereotypic Explanatory Bias: Implicit Stereotyping as a Predictor of Discrimination, 39 J. Experimental Soc. Psychol.75 (2003).

Dolly Chugh, Societal and Managerial Implications of Implicit Social Cognition: Why Milliseconds Matter, 17 Soc. Just. Res. 203 (2004)

Elizabeth A. Phelps, et. al., Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation, 12 J. Cog. Neuroscience 729 (2000).

Eric Luis Uhlmann & Geoffrey L. Cohen, Constructed Criteria: Redefining Merit to Justify Discrimination, 16 Psychol. Sci. 474 (2005)

Eugene M. Caruso, Dobromir A. Rahnev, & Mahzarin Banaji, Using Conjoint Analysis to Detect Discrimination: Revealing Covert Preferences from Overt Choices, 27 Soc. Cognition 128 (2009)

Gregory Mitchell & Philip E. Tetlock, Antidiscrimination Law and the Perils of Mindreading, 67 Ohio St. L.J. 1023 (2006)

Hal R. Arkes & Philip E. Tetlock, Attributions of Implicit Prejudice, or "Would Jesse Jackson 'Fail' the Implicit Association Test?," 15 Psychol. Inquiry 257 (2004)

Irene V. Blair et al, The Influence of Afrocentric Facial Features in Criminal Sentencing, 15 Psychol. Sci. 674 (2004)

Irene V. Blair, Jennifer E. Ma, & Alison P. Lenton, Imagining Stereotypes Away: The Moderation of Implicit Stereotypes Through Mental Imagery, 81 J. Person. & Soc. Psych. 828 (2001)

Jerry Kang, Trojan Horses of Race, 118 Harv. L. Rev. 1491 (2005)

Kurt Hugenberg & Galen V. Bodenhausen, Facing Prejudice: Implicit Prejudice and the Perception of Facial Threat. 14 Psychol. Sci. 640 (2003)

Laurie A. Rudman & Matthew R. Lee, Implicit and Explicit Consequences of Exposure to Violent and Misogynous Rap Music, 5 Group Processes & Intergroup Rel. 133 (2002)

Laurie A. Rudman & Peter Glick, Prescriptive Gender Stereotypes and Backlash Toward Agentic Women, 57 J. Soc. Issues 743 (2001)

Luciano Arcuri et al., Predicting the Vote: Implicit Attitudes as Predictors of the Future Behavior of Decided and Undecided Voters, 29 Political Psychology369

Matthew K. Knock & Mahzarin R. Banaji, Prediction of Suicide Ideation and Attempts Among Adolescents Using a Brief Performance-Based Test, 75 J. Clinical & Consulting Psychology 707 (2007).

Nicola S. Gray, et al., An Implicit Test of the Associations Between Children and Sex in Pedophiles, 114 J. Abnormal Psych. 304 (2005)

Nilanjana Dasgupta & Anthony G. Greenwald, On the Malleability of Automatic Attitudes: Combating Automatic Prejudice with Images of Admired and the Disliked Individuals, 81 J. Person. & Soc. Psych.800 (2001)

Nilanjana Dasgupta & Shaki Asgari, Seeing is Believing: Exposure to Counterstereotypic Women Leaders and Its Effect on the Malleability of Automatic Gender Stereotyping, 40 J. Experimental Soc. Psychol. 642 (2004)

R. Richard Banks, Jennifer L. Eberhardt, Lee Ross, Discrimination and Implicit Racial Bias in a Racially Unequal Society. 94 Calif. Law Rev. 1169 (2006)

William Von Hippel et al., The Linguistic Intergroup Bias As an Implicit Indicator of Prejudice, 33 J. Experimental Soc. Psychol. 490 (1997)