CONNECTICUT
LAW REVIEW

VOLUME 58 | FEBRUARY 2026 | NUMBER 2

Article

Misjudging a Reasonable Jury: Evidence That Courts Dismiss Meritorious Harassment Claims

ELIZABETH C. TIPPETT & JAMILLAH BOWMAN WILLIAMS

Courts assessing summary judgment motions in Title VII harassment claims commonly grant the motion on the basis that the alleged harassment is insufficiently "severe or pervasive" to meet the legal standard. This mixed-methods study empirically tests whether there is a gap between how judges and potential jurors assess the same set of facts on the severe or pervasive element of a Title VII harassment claim. We presented study participants with facts from 80 federal harassment cases. In each case, the defendant employer moved to dismiss the case, arguing that no reasonable jury would find the alleged harassment sufficiently severe or pervasive to meet the legal threshold. We provided the participants with relevant jury instructions and asked them to: (1) rate the severity or pervasiveness of the alleged harassment; (2) assess whether the plaintiff met the legal standard; and (3) discuss their reasoning.

Our results suggest a substantial divergence between judicial assessments and simulated jury assessments of the sampled cases. Judges granted summary judgment in favor of the employer or dismissed 65% of the harassment cases in the sample. By contrast, our simulated juries would have dismissed less than 20% of the very same cases. Both our quantitative and qualitative findings shed light on the source of this divergence. The difference in assessment is not due to demographic differences between judges and the mock jurors, nor is it caused by shifting judicial assessments over time. Our qualitative results indicate laypeople tend to view the fact patterns in a much more holistic manner than judges, which is consistent with guidance established by the Supreme Court. In addition, our quantitative analysis suggests courts may be selectively discounting the severity or pervasiveness of cases alleging intersectional harassment based on more than one protected characteristic. Going forward, we recommend that courts exercise far greater caution in evaluating harassment claims on summary judgment. Courts should also allow intersectional claims to be pled as a single cause of action.

ARTICLE CONTENTS

Misjudging a Reasonable Jury: Evidence That Courts Dismiss Meritorious Harassment Claims

ELIABETH C. TIPPETT* & JAMILLAH BOWMAN WILLIAMS**

INTRODUCTION

Several studies have documented the unusually high rates of summary judgment in employment discrimination claims—including harassment claims—under Title VII of the Civil Rights Act, which prohibits discrimination on the basis of race, color, religion, sex, or national origin.[1] Employment law scholars also critique individual lower court rulings for granting summary judgment in discrimination and harassment claims despite strong evidence in support of the plaintiff's claims.[2]

This study is the first to bridge these critical intuitions of employment scholars with the empirical literature that has long demonstrated high rates of summary judgment in favor of employers. We utilized an innovative mixed-methods research design, which presented the facts and jury instructions for 80 real Title VII harassment claims to laypeople. We focused specifically on cases where the "severe or pervasive" element of harassment claims was at issue on summary judgment, as scholars have consistently critiqued the manner in which judges have applied this particular standard in their rulings. Generally, the severe or pervasive standard assesses the extent to which harassment is both objectively and subjectively hostile or abusive.[3] In lay terms, one may say this standard measures whether the harassing conduct was bad or objectionable enough to give rise to a legal claim.

Our study included both quantitative and qualitative data. We asked our 699 respondents to (1) rate the severity or pervasiveness of the fact pattern from a single case on a scale of 0–100, (2) assess whether the plaintiff met

---

* Professor and James O. and Alfred T. Goodwin Senior Fellow, University of Oregon School of Law. We are grateful to Yi Yao, Meg Kirschnick, Brett Henson, Catherine Roner Reiter, Georgia Zacest, Hayley Cloyd, Anu Ramdin, and Edriana Jones for research support. Thank you to Jane Larrington, Scott Dodson, Nicole Porter, Orli Oren Kolbinger, Michael Moffitt, Erik Girvan, Mindy Wittkop, and Jon Lee for their generous feedback and commentary at various stages of the research process.

** Professor of Law and Anne Fleming Research Professor, Georgetown University Law Center, Faculty Affiliate, Georgetown Department of Sociology, and Founding Faculty Director, Georgetown Workers' Rights Institute.

[1] Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e to 2000e-17.

[2] See *infra* Part I.D for a discussion regarding how courts discredit strong evidence and create bad precedent.

[3] See *infra* Part I.B for an explanation of the elements of a Title VII harassment claim, including the "severe or pervasive" element.

the legal standard for severe or pervasive harassment, and (3) explain the reasoning behind their conclusion. We also collected demographic information that enabled us to assess whether demographic differences between judges and juries could account for any difference in assessments and decision-making.

Because we used fact patterns from real cases, we were able to compare the respondents' assessments with the outcome of the case. We could therefore analyze empirically whether judges had correctly applied the summary judgment standard under Rule 56 of the Federal Rules of Civil Procedure: "whether a fair-minded jury could return a verdict for the plaintiff on the evidence presented."[4] We also collected case-level information regarding the plaintiff's protected category, the employer type (government or private sector), and the year the case was decided. This allowed us to examine whether these factors predict any observed disparities between courts and simulated jurors.

Broadly, we found that courts are not accurately forecasting how jurors would assess plaintiffs' harassment claims and are far too aggressive in dismissing cases on the severe or pervasive element. In 83% of cases in the sample, at least half of the mock jurors concluded the severe or pervasive standard was met. For the purposes of this study, we used this 50% threshold to approximate how a "fair-minded" jury could rule.[5] By contrast, a much smaller proportion of cases (35%) actually survived summary judgment or a motion to dismiss and had an opportunity to be considered by a jury of the plaintiff's peers. Our quantitative analyses test a variety of hypotheses to explain this divergence between judicial and layperson assessments, which we describe in greater detail in Part IV below.

Our qualitative analysis, set forth in Part V below, provides a much more detailed account of the differences between judges and juries derived from the narrative responses of our mock jury respondents. Our comparison of the respondents' reasoning with the court decisions suggests that respondents engaged in a much more holistic and contextual assessment of the evidence, considering a variety of factors that could affect plaintiffs' experience of the workplace. These factors included those that the Supreme Court has enumerated as legally significant, such as the frequency of the harassment and whether there was physical contact.[6] However, our respondents also took other facts into account, such as the number of harassers, the number of victims, whether the harassment created obstacles to performing the plaintiff's job, and whether the harasser or the organization ignored complaints. All these factors might render a workplace more hostile to an average worker.

---

[4] Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 252 (1986); *see also* FED. R. CIV. P. 56(a).
[5] *Anderson*, 477 U.S. at 252.
[6] Harris v. Forklift Sys., Inc., 510 U.S. 17, 23 (1993).

Although judges sometimes considered these aggravating factors in their assessment of whether the conduct qualified as severe or pervasive, many others overlooked or discounted them. Some courts seemed afflicted by a kind of "tunnel vision" or—what legal scholars have alternatively characterized as a "slice and dice"[7] or "divide and conquer" approach[8]—where courts allocate evidence to other elements of the claim or to other claims entirely and fail to view the circumstances as a whole. In a way, their legal expertise, which trains them to break down and sort evidence based on legal elements, hampered their ability to assess how a layperson might experience a hostile workplace in totality.

These results lead to two principal policy recommendations, discussed in Part V. First, courts should be far more reticent to grant summary judgment in harassment claims on the basis that no reasonable jury would find in the plaintiff's favor on the severe or pervasive element. Second, the Equal Employment Opportunity Commission (EEOC) and courts should interpret Title VII to permit "intersectional" claims—those based on more than one category protected by the statute (e.g., race and sex)—to be asserted as a single cause of action, rather than multiple separate causes of action for each protected class.[9] Doing so would remove the heightened evidentiary burdens currently required of plaintiffs experiencing intersectional harassment.

The Article proceeds as follows: Part I provides the legal context for summary judgment generally, and the harassment context specifically, followed by a discussion of relevant empirical studies and scholarly debates. Part II describes our methodology and the research questions that guide the study. Part III discusses the patterns that emerge in our quantitative findings. Part IV describes our qualitative findings, which offer deeper insights into each mock juror's decision-making process. Part V considers the implications of this research project, followed by a Conclusion.

I. LEGAL CONTEXT

The following Section presents the legal and scholarly context that guides this research, beginning with an overview of the legal standard for granting summary judgment. We discuss the leading Supreme Court cases that operationalize the summary judgment standard and discuss related

[7] Michael J. Zimmer, *Slicing & Dicing of Individual Disparate Treatment Law*, 61 LA. L. REV. 577, 577 (2001).

[8] Theresa M. Beiner, *Let the Jury Decide: The Gap Between What Judges and Reasonable People Believe Is Sexually Harassing*, 75 S. CAL. L. REV. 791, 808 (2002).

[9] *See* Patrick Berning-O'Neill, *"A Reasonably Comparable Evil": Expanding Intersectional Claims Under Title VII Using Existing Precedent*, 24 U. PA. J. CONST. L. 907, 908 (2022) (arguing that intersectional discrimination claims are "a reasonably comparable evil" to single-basis discrimination and thus should fall under Title VII).

scholarly commentary. Next, we summarize the legal standard for Title VII harassment claims. We then review existing literature that provides empirical evidence that plaintiffs in harassment claims fare poorly on summary judgment relative to other types of claims. Finally, we engage with scholarly commentary hypothesizing that misapplication of the "severe or pervasive" standard leads to these poor outcomes for Title VII plaintiffs.

A. *The Legal Standard for Granting Summary Judgment*

Summary judgment motions enable a party to move for the dismissal of a case prior to trial, often after some or all discovery is complete.[10] As a practical matter, summary judgment rulings are almost always filed by defendants, because the procedural posture of the motion favors the defendant.[11] On summary judgment, a defendant need only show that one element of the claim is fatally defective—that is, an element for which there is no disputed issue of fact—for that entire cause of action to be dismissed.[12] By contrast, for a plaintiff to prevail on summary judgement, they must prove that there is no disputed issue of fact as to every element of the claim.[13] Consequently, summary judgment rulings overwhelmingly represent an opportunity for the court to rule in the defendant's favor and prevent the case from going to trial.

The standard for summary judgment is set forth in Federal Rule of Civil Procedure 56, which instructs the court to grant summary judgment "if the movant shows that there is no genuine dispute as to any material fact and the movant is entitled to judgment as a matter of law."[14] The question, however, of what qualifies as a "genuine dispute as to any material fact" has been a matter of significant controversy. Prior to 1986, courts treated summary judgment as equivalent to a Rule 12(b)(6) motion. Under that framework, courts would only grant summary judgment if the case was not "cognizable" even after discovery was complete—that is, "only when the law offered no possible redress for any of the facts in the full pretrial record."[15] According to commentator Steven Childress, "[d]istrict courts were warned incessantly not to weigh evidence or to resolve factual disputes."[16] The prevailing standard among circuit courts in the 1970s was one where summary judgment should

---

[10] FED. R. CIV. P. 56(b) (authorizing the filing of a motion for summary judgment "at any time until 30 days after the close of all discovery").

[11] Elizabeth C. Tippett, Charlotte S. Alexander, Karl Branting, Paul Morawski, Carlos Balhana, Craig Pfeifer & Sam Bayer, *Does Lawyering Matter? Predicting Judicial Decisions from Legal Briefs, and What That Means for Access to Justice*, 100 TEX. L. REV. 1157, 1168–69 (2022) (asserting that the defendant has "a much easier burden when moving for summary judgment").

[12] *Id.*

[13] *Id.* at 1168.

[14] FED. R. CIV. P. 56(a).

[15] Steven Alan Childress, *A New Era for Summary Judgments: Recent Shifts at the Supreme Court*, 6 REV. LITIG. 263, 264 (1987).

[16] *Id.* at 264–65.

be denied "if there is 'the slightest doubt' as to its propriety."[17] Commentator Paul Mollica observed that cases from the 1970s evinced an "extreme vigilance against treading on contested fact issues or mixed questions of law and fact—even arguable ones—reserving them for evidentiary hearings."[18]

Three Supreme Court cases decided in 1986—*Matsushita Electric Industries Co. v. Zenith Radio Corp.*,[19] *Celotex Corp. v. Catrett*,[20] and *Anderson v. Liberty Lobby, Inc.*[21]—transformed how lower courts approached summary judgment.[22] In *Matsushita*, the Supreme Court articulated the summary judgment standard commonly applied today, that "[w]here the record taken as a whole could not lead a rational trier of fact to find for the non-moving party, there is no 'genuine issue for trial.'"[23] In *Celotex*, the Court asserted that summary judgment should not be treated as a "disfavored procedural shortcut, but rather as an integral part of the Federal Rules . . . ."[24] Instead, it should serve to balance the right to a jury with the "rights of persons opposing such claims and defenses" to demonstrate "that the claims and defenses have no factual basis."[25]

Both *Liberty Lobby* and *Matsushita* characterized summary judgment as akin to a Rule 50 motion for a directed verdict. Childress argued that this comparison was a significant shift from prior jurisprudence because it authorized trial courts to weigh the sufficiency of the evidence and thereby "weed[] out" cases with "weak factual claims."[26] As the Court instructed in *Liberty Lobby*:

> [T]he judge must ask himself . . . whether a fair-minded jury could return a verdict for the plaintiff on the evidence presented. The mere existence of a scintilla of evidence in

---

[17] *Id.* at 264 (quoting Doehler Metal Furniture Co. v. United States, 149 F.2d 130, 135 (2d Cir. 1945)). *See generally* Michael E. Smith, *Judge Charles E. Clark and the Federal Rules of Civil Procedure*, 85 YALE L.J. 914, 928–31 (1976) (discussing summary judgment jurisprudence at that time).

[18] Paul W. Mollica, *Federal Summary Judgment at High Tide*, 84 MARQ. L. REV. 141, 147 (2000).

[19] Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574, 587 (1986).

[20] Celotex Corp. v. Catrett, 477 U.S. 317, 327 (1986).

[21] Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 256 (1986).

[22] *See* Childress, *supra* note 15, at 265 (describing the "doctrinal development" that "increases the situations in which a grant would be appropriate"). *See, e.g.*, Suja A. Thomas, *The Fallacy of Dispositive Procedure*, 50 B.C. L. REV. 759, 765 (2009) (discussing the impact of *Anderson* on the summary judgment standard); Henry L. Chambers, Jr., *Recapturing Summary Adjudication Principles in Disparate Treatment Cases*, 58 SMU L. REV. 103, 107 (2005) (discussing the impact of the trilogy cases on the summary judgment standard); Mollica, *supra* note 18, at 164 (same); Michael S. Pardo, *Pleadings, Proof, and Judgment: A Unified Theory of Civil Litigation*, 51 B.C. L. REV. 1451, 1501 (2010) (same).

[23] *Matsushita*, 475 U.S. at 587 (quoting First Nat'l Bank of Ariz. v. Cities Serv. Co., 391 U.S. 253, 288 (1968) (citation omitted)).

[24] *Celotex*, 477 U.S. at 327; *see also* Childress, *supra* note 15, at 272 (discussing the Court's use of broad language favoring summary judgment in *Matsushita* and *Celotex*).

[25] *Celotex*, 477 U.S. at 327.

[26] Childress, *supra* note 15, at 267.

> support of the plaintiff's position will be insufficient; there must be evidence on which the jury could reasonably find for the plaintiff.[27]

In 1987, Childress predicted that *Matsushita*, *Celotex*, and *Liberty Lobby* would open the door to "a new and liberal standard of granting motions for summary judgment."[28] His prediction proved correct. The three 1986 cases are today referred to as the "trilogy" that remain the leading Supreme Court cases on summary judgment.[29] Although the trilogy cases did not formally alter the legal standard, they appear to have altered decision-making by lower courts. Mollica compared summary judgment rulings in the 1970s and 1990s and concluded that "the variety of cases and situations in which federal courts granted summary judgment widened over the study period."[30]

A fourth Supreme Court case decided in 2007, *Scott v. Harris*, has also generated significant scholarly attention.[31] *Scott v. Harris* took an even more expansive view of summary judgment than the trilogy cases, declaring that the "facts must be viewed in the light most favorable to the nonmoving party only if there is a 'genuine' dispute as to those facts."[32] In *Scott v. Harris*, the plaintiff brought a Section 1983 case against a police officer for using excessive force by ramming into the plaintiff's car following a high-speed chase.[33] The evidence included a video of the chase, which the justices apparently watched. Based on the video, the Court concluded that the

---

[27] Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 252 (1986).

[28] Childress, *supra* note 15, at 265.

[29] Thomas, *supra* note 22, at 765; Chambers, Jr., *supra* note 22, at 107; Mollica, *supra* note 18, at 164. The 1986 trilogy cases remain the most cited cases on summary judgment. Citation metrics on Google Scholar suggest that *Liberty Lobby* is the most cited case on point with more than 230,000 citations. Search results for "Anderson v. Liberty Lobby, Inc.," GOOGLE SCHOLAR, https://scholar.google.com/scholar?hl=en&as_sdt=8006&q=Anderson+v.+Liberty+Lobby%2C+Inc.&btnG= (last visited Oct. 24, 2025). *Celotex* follows, with more than 210,000 citations. Search results for "Celotex Corp. v. Catrett," GOOGLE SCHOLAR, https://scholar.google.com/scholar?hl=en&as_sdt=8006&q=Celotex+Corp.+v.+Catrett&btnG= (last visited Oct. 24, 2025). Finally, *Matsushita* trails with more than 110,000 citations. Search results for "Matsushita Elec. Indus. Co. v. Zenith Radio Corp.," GOOGLE SCHOLAR, https://scholar.google.com/scholar?hl=en&as_sdt=8006&q=Matsushita+Elec.+Indus.+Co.+v.+Zenith+Radio+Corp.&btnG= (last visited Oct. 24, 2025).

[30] Mollica, *supra* note 18, at 144.

[31] Scott v. Harris, 550 U.S. 372, 380 (2007). We do not discuss *Reeves v. Sanderson Plumbing Prods., Inc.* in this Section as that case is generally understood among employment law scholars as a substantive interpretation of the *McDonnell Douglas* standard under the Age Discrimination in Employment Act—and ultimately Title VII of the Civil Rights Act—rather than a procedural ruling. Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 148 (2000) (stating that causation in ADEA cases can be proven through pretext alone); McDonnell Douglas Corp v. Green, 411 U.S. 792, 802 (1973) (articulating the prima facie case in a Title VII discrimination case).

[32] *Scott*, 550 U.S. at 380. Commentator Tobias Wolff observed that the case introduced some doctrinal instability into lower court jurisprudence because the language suggested a departure from longstanding jurisprudence, where the nonmoving party is entitled to deference throughout summary judgment process. Tobias Barrington Wolff, Scott v. Harris *and the Future of Summary Judgment*, 15 NEV. L.J. 1351, 1358 (2015). However, the court has since backed away from that stance in recent jurisprudence, placing *Scott v. Harris* more in line with the 1986 trilogy. *Id.* at 1353.

[33] *Scott*, 550 U.S. at 375–76.

plaintiff's factual assertions—including that he was not "a threat to pedestrians"—were "utterly discredited by the record" such that "no reasonable jury could have believed him."[34]

Commentator Suja Thomas critiqued the Court in *Scott v. Harris* for substituting its own judgment for the jury.[35] More broadly, Thomas argued "the reasonable jury standard has become a proxy for a judge's own view of the evidence,"[36] a critique that will prove to be particularly relevant in the harassment cases we examine. Similarly, Dan Kahan et al. offered an empirically informed critique of *Scott v. Harris*, showing the video footage from the case to a sample of 1,350 Americans and assessing whether demographic differences and cultural attitudes altered their interpretation.[37] Kahan found that Black respondents and low-income respondents found the use of deadly force to be less justifiable.[38] Kahan characterizes the problem as one of "cognitive illiberalism"—where judges fail to consider how their own views of the facts might not be shared by others.[39]

Notwithstanding scholarly critique of the 1986 trilogy and the *Scott v. Harris* case, the legal standard for summary judgment remains quite stringent and should be reserved for cases where no "fair-minded jury could return a verdict for the plaintiff on the evidence presented."[40] Even so, rates of summary judgment are quite high in Title VII discrimination cases—a category that includes harassment cases. Next, we summarize the legal standard for Title VII harassment cases, as well as empirical research and scholarly commentary regarding summary judgment in the harassment context.

B. *The Legal Standard for Harassment Under Title VII of the Civil Rights Act*

Title VII of the Civil Rights Act of 1964 prohibits employment discrimination on the basis of "race, color, religion, sex, or national origin .

---

[34] *Id.* at 379–80 (citation omitted).

[35] Thomas, *supra* note 22, at 770. Thomas notes that the interpretation of the video could not actually have been uniform if lower courts and dissenting justices disagreed. *Id.* at 772–73.

[36] Suja A. Thomas, *Summary Judgment and the Reasonable Jury Standard: A Proxy for a Judge's Own View of the Sufficiency of the Evidence?*, 97 JUDICATURE 222, 223 (2014). Thomas also observes that courts tend to use the terms "reasonable jury," "reasonable juror," "rational juror," and "rational factfinder" somewhat "interchangeably" even though the different formulations would likely produce different outcomes. Thomas, *supra* note 22, at 771–72.

[37] Dan M. Kahan, David A. Hoffman & Donald Braman, *Whose Eyes are you Going to Believe?* Scott v. Harris *and the Perils of Cognitive Illiberalism*, 122 HARV. L. REV. 837, 854 (2009).

[38] *Id.* at 841, 867.

[39] *Id.* at 895–96. Kahan's study also inadvertently illustrated the unusual character of the evidence of the *Scott v. Harris* case. Despite the between-group differences Kahan described, a very large majority of respondents (75%) concluded that deadly force was justified after watching the video. *Id.* at 866. While such figures do not represent unanimity, the empirical results lend some support to the Court's assertion that the video rendered the evidence heavily lopsided in favor of the defendant.

[40] Anderson v. Liberty Lobby, Inc., 477 U.S. 242, 252 (1986).

. . .”[41] The Supreme Court first recognized harassment as a form of prohibited Title VII discrimination in the 1986 case *Meritor Savings Bank v. Vinson*. In that case, the Court declared that “a plaintiff may establish a violation of Title VII” where the sex-based harassment is “sufficiently severe or pervasive to alter the conditions of . . . employment and create an abusive working environment.”[42]

The Supreme Court further defined the contours of Title VII harassment claims in a series of decisions in the 1990s.[43] Two of these opinions are particularly relevant for our purposes: *Harris v. Forklift Sys., Inc.*, and *Oncale v. Sundowner Offshore Services, Inc*.[44] *Harris v. Forklift* elaborated on the severe or pervasive standard, stating that harassment must be both objectively and subjectively hostile or abusive, meaning that “a reasonable person would find [the environment] hostile or abusive”[45] and that “the victim . . . subjectively perceive[s] the environment to be abusive . . . .”[46] Assessments of the severity or pervasiveness of the conduct should consider “all the circumstances,” including:

> the frequency of the discriminatory conduct; its severity; whether it is physically threatening or humiliating, or a mere offensive utterance; and whether it unreasonably interferes with an employee’s work performance.[47]

In *Oncale v. Sundowner*, the Court recognized same-sex harassment as a form of harassment. Although the case did not implicate the severe or pervasive standard, the Court observed that the severe or pervasive requirement distinguished prohibited harassment from a “general civility code” or “ordinary socializing in the workplace—such as male-on-male horseplay or intersexual flirtation . . . .”[48]

Current jury instructions in Title VII harassment claims reflect a blend of the language in *Meritor*, *Harris v. Forklift*, and *Oncale*. Federal circuits vary somewhat in the precise wording of jury instructions, with alternate formulations depending on whether the harassment included a tangible

---

[41] 42 U.S.C. § 2000e-2(a)(1).

[42] Meritor Sav. Bank, FSB v. Vinson, 477 U.S. 57, 66–67 (1986) (internal quotation marks omitted) (citation omitted).

[43] Harris v. Forklift Sys., Inc., 510 U.S. 17, 21 (1993); Oncale v. Sundowner Offshore Servs., 523 U.S. 75, 80–81 (1998); Faragher v. City of Boca Raton, 524 U.S. 775, 780 (1998) (articulating circumstances in which employers will be held vicariously liable for harassment by employees); Burlington Indus., Inc., v. Ellerth, 524 U.S. 742, 765 (1998) (same). Harassment jurisprudence in recent decades has not significantly altered the elements of Title VII harassment claims. *See* Vance v. Ball State Univ., 570 U.S. 421, 431 (2013) (defining “supervisor” for the purpose of vicarious liability); Bostock v. Clayton Cnty., 140 S. Ct. 1731, 1742 (2020) (stating that sexual orientation and gender identity discrimination qualify as discrimination on the basis of “sex”).

[44] *Harris*, 510 U.S. at 17; *Oncale*, 523 U.S. at 75.

[45] *Harris*, 510 U.S. at 21.

[46] *Id.* To be subjectively abusive, the environment need not necessarily be “psychologically injurious” as long as it is “perceived . . . as hostile or abusive.” *Id.* at 22.

[47] *Id.* at 23.

[48] *Oncale*, 523 U.S. at 81.

employment action and whether the harassment was committed by a supervisor. However, the instructions generally provide that a harassment plaintiff must prove the following five elements:

> 1. First, the plaintiff was subjected to [describe alleged conduct or conditions giving rise to the plaintiff's claim]; and
> 2. Second, such conduct was unwelcome; and
> 3. Third, such conduct was based on the plaintiff's [(sex/gender) (race) (color) (national origin) (religion)]; and
> 4. Fourth, such conduct was sufficiently severe or pervasive that a reasonable person in the plaintiff's position would find the plaintiff's work environment to be [(hostile) (abusive)]; and
> 5. Fifth, at the time such conduct occurred and as a result of such conduct, the plaintiff believed [(his) (her)] work environment to be [(hostile) (abusive)].[49]

C. *Research Suggests Title VII Harassment Plaintiffs Fare Poorly on Summary Judgment*

Several legal scholars have critiqued the high rates of summary judgment in Title VII cases generally, and for harassment claims specifically. Sandra Sperino analyzed a variety of harassment cases where courts have granted summary judgment notwithstanding "evidence that the conditions of employment have been affected by sex or race."[50] Tristin

---

[49] MANUAL OF MODEL CIV. JURY INSTRUCTIONS FOR THE DIST. CTS. OF THE 8TH CIR., Instruction 8.41: Harassment (By Supervisor With No Tangible Employment Action) (Comm. on Model Jury Instructions for The Dist. Courts of the 8th Cir., amended 2018). We cite older versions of model jury instructions because the instructions herein correspond to those in effect when the study was designed and data was initially collected. *See also* 3D CIR. MODEL CIV. JURY INSTRUCTIONS, Instruction 5.1.5: Elements of a Title VII Claim—Harassment—Hostile Work Environment—No Tangible Employment Action (Comm. on Model Civil Jury Instructions Within the 3d Cir., amended 2018) (nearly identical to Eighth Circuit instructions, but includes an additional sentence elaborating on the objectively hostile component stating, "This element requires you to look at the evidence from the point of view of a reasonable [member of plaintiff's protected class] reaction to [plaintiff's] work environment"); MANUAL OF MODEL CIV. JURY INSTRUCTIONS FOR THE DIST. CTS. OF THE 9TH CIR., Instruction 10.2A: Civil Rights—Title VII—Hostile Work Environment—Harassment Because of Protected Characteristics—Elements (9th Cir. Jury Instructions Comm., amended 2016), (very similar to Eighth Circuit formulation); FED. CIV. JURY INSTRUCTIONS OF THE 7TH CIR., 3.05B Supervisor Harassment with No Tangible Employment Action (Comm. on Pattern Civil Jury Instructions of the 7th Cir., amended 2015). *Cf.* PATTERN JURY INSTRUCTIONS (CIV. CASES), Instruction 11.2 (Comm. on Pattern Jury Instructions Dist. Judges Ass'n 5th Cir., amended 2016) (presenting the instructions in paragraph form, rather than a numbered list, and including the language in *Harris v. Forklift* regarding the frequency of the conduct and other factors).

[50] Sandra F. Sperino, *Rethinking Discrimination Law*, 110 MICH. L. REV. 69, 87 (2011). Sperino

Green likewise identified a series of egregious race-based harassment cases that were dismissed on summary judgment as insufficiently severe or pervasive.[51]

Statistics suggest that summary judgment has been granted at high rates in employment discrimination cases, which include harassment claims. Broadly, discrimination claims fare more poorly than other civil cases. Clermont and Schwab's study of federal employment cases from 1979 to 2006 estimated a win rate of 15% for employment discrimination plaintiffs, compared to 51% for all other cases.[52] In addition, a 2005 PACER based study by Vivian Berger et al. found that some 70% of employment discrimination cases do not survive summary judgment.[53] Likewise, a 2012 study of summary judgment rulings in employment civil rights cases found that 56% of filings resulted in dismissal of the entire case.[54]

Researchers have also conducted "policy-capturing" studies of harassment cases, a quantitative methodology that attempts to identify case characteristics that are most predictive of legal outcomes.[55] A 2001 study of summary judgment rulings in sexual harassment cases found that the frequency and severity of the conduct, and the presence of witnesses, were statistically predictive of the case outcome.[56] A study by Ann Juliano and Stewart J. Schwab found that plaintiffs were more likely to prevail when the harassment included physical contact.[57] They were also more successful when the conduct involved "belittling or derogatory comments," comments

---

also argues that the severe or pervasive requirement is unsupported by the statutory text, which only requires that the conduct "affect[s] the terms or conditions of . . . plaintiff's employment . . . ." *Id.* at 97; *see also* Sandra F. Sperino & Suja A. Thomas, *Fakers and Floodgates*, 10 STAN. J. C.R. & C.L. 223, 248 (discussing evidence courts deemed insufficient to meet the legal threshold for harassment); SANDRA F. SPERINO & SUJA A. THOMAS, UNEQUAL: HOW AMERICA'S COURTS UNDERMINE DISCRIMINATION LAW 19–23, 32–37 (2017) (examining cases where judges have discounted evidence of harassment); Beiner, *supra* note 8, at 842 (discussing the gap between what laymen consider harassment and what judges have found to legally constitute harassment).

[51] TRISTIN K. GREEN, RACIAL EMOTION AT WORK: DISMANTLING DISCRIMINATION AND BUILDING RACIAL JUSTICE IN THE WORKPLACE 69–70 (2023).

[52] Kevin M. Clermont & Stewart J. Schwab, *Employment Discrimination Plaintiffs in Federal Court: From Bad to Worse?*, 3 HARV. L. & POL'Y REV. 103, 127 (2009).

[53] Vivian Berger, Michael O. Finkelstein & Kenneth Cheung, *Summary Judgment Benchmarks for Settling Employment Discrimination Lawsuits*, 23 HOFSTRA LAB. & EMP. L.J. 45, 55 (2005); *see also* Deborah Erdos Knapp & Brian P. Heshizer, *Outcomes of Requests for Summary Judgments in Federal Sexual Harassment Cases: Policy Capturing Revisited*, 44 SEX ROLES 109, 117 (2001) (analyzing the outcomes of early harassment claims from the late 1990s).

[54] Jill D. Weinberg & Laura Beth Nielsen, *Examining Empathy: Discrimination, Experience, and Judicial Decisionmaking*, 85 S. CAL. L. REV. 313, 330–31, 336 (2012).

[55] Mark V. Roehling, *"Extracting" Policy from Judicial Opinions: The Dangers of Policy Capturing in a Field Setting*, 46 PERS. PSYCH. 477, 478 (1993); Knapp & Heshizer, *supra* note 53, at 112.

[56] Knapp & Heshizer, *supra* note 55, at 123 (finding that five of the measured variables had a significant effect on the outcome of the case); Ann Juliano & Stewart J. Schwab, *The Sweep of Sexual Harassment Cases*, 86 CORNELL L. REV. 548, 572 (2001) (Juliano and Schwab likewise found that plaintiffs were more likely to prevail if the "[p]laintiff told coworkers of harassment").

[57] Juliano & Schwab, *supra* note 56, at 571.

about the plaintiff's "physical appearance," or when "comments of a sexual nature" were directed "specifically to" the plaintiff.[58]

Most of the studies and commentary regarding the severe or pervasive requirement were specific to sexual harassment claims and did not address harassment based on other protected categories. However, broader studies on employment discrimination, including harassment, note that some types of claims fare better than others. For example, Berger's study of summary judgment in employment discrimination claims found that "sexual harassment" claims fared better on summary judgment than all other types of discrimination claims.[59] Berger's study, however, involved only 21 sexual harassment cases, which limits the generalizability of the findings.[60]

High rates of summary judgment in discrimination and harassment claims are concerning if the plaintiffs would have prevailed had their case proceeded before a jury. There is some limited evidence that juries would view the facts more favorably than courts. For example, plaintiffs fare better in jury trials than bench trials in employment discrimination cases. A study by Kevin Clermont and Stewart J. Schwab found that plaintiffs prevailed in approximately 38% of jury trials, compared to 19% of bench trials.[61] A variety of scholars have argued that the demographics and life experiences of judges might affect their attitudes towards employment discrimination claims.[62] Some empirical research tentatively supports this proposition. Jill D. Weinberg and Laura Beth Nielson's study found that judges of color granted summary judgment in employment cases at much lower rates than white judges (38% versus 61%).[63] In addition, White judges were more favorably disposed to claims brought by White plaintiffs: "white judges tend to dismiss cases involving minority plaintiffs at a much higher rate than cases involving white plaintiffs."[64] Weinberg and Nielson conclude that such

---

[58] *Id.*

[59] Berger, Finkelstein & Chung, *supra* note 53, at 60.

[60] *Id.*

[61] Kevin M. Clermont & Stewart J. Schwab, *How Employment Discrimination Plaintiffs Fare in Federal Court*, 1 J. EMPIRICAL LEGAL STUD. 429, 442 (2004). *Cf.* Tamar Kricheli-Katz & Keren Weinshall, *Judging Fast or Slow: The Effects of Reduced Caseloads on Gender- and Ethnic-Based Disparities in Case Outcomes*, 20 J. EMPIRICAL LEGAL STUD. 961, 963 (2023) (finding that judges in Israel were less biased against female and Arab plaintiffs when they had reduced caseloads).

[62] *See, e.g.*, Beiner, *supra* note 8, at 820–21 n.146 (arguing that federal judges "may not be in the best position" to assess whether conduct is severe or pervasive given their gender and income, and that "their ability to empathize with the common worker may be reasonably questioned"); Theresa M. Beiner, *The Misuse of Summary Judgment in Hostile Environment Cases*, 34 WAKE FOREST L. REV. 71, 122–27 (1999) [hereinafter *Misuse of Summary Judgment*]; Leslie M. Kerns, *A Feminist Perspective: Why Feminists Should Give the Reasonable Woman Standard Another Chance*, 10 COLUM. J. GENDER & L. 195, 210 (2001) ("[B]ecause the majority of judges are men who have likely experienced, or have been leaders in, a male-dominated workplace, it is natural for male judges to accept the perspective of the male aggressor over the female victim.").

[63] Jill D. Weinberg & Laura Beth Nielsen, *Examining Empathy: Discrimination, Experience, and Judicial Decisionmaking*, 85 S. CAL. L. REV. 313, 339 (2012).

[64] *Id.* at 346.

outcomes result from "different attitudes, opinions, and experiences that stem from being white or a person of color."[65]

Judges, however, do not have a monopoly on bias. Even if juries respond more favorably than courts, their verdicts may likewise reflect more favorable attitudes to some types of plaintiffs than others. Oppenheimer's examination of jury verdicts in California employment law cases found that sexual harassment cases fared better than other types of discrimination cases, prevailing in 67% of cases, and "with a median verdict of $210,000."[66] Race-based harassment claims brought by people of color prevailed in about 50% of cases.[67] Discrimination cases brought by Black women had the lowest success rate of all discrimination cases at 17%.[68] Noting the particularly poor outcomes for race discrimination cases—and Black women specifically—Oppenheimer concludes that "juror bias and judicial bias are both likely suspects."[69]

Oppenheimer's results support Kimberle Crenshaw's theory of intersectional subordination.[70] In 1989, Crenshaw coined the term to highlight how anti-discrimination laws and policies fail to acknowledge compounded forms of discrimination.[71] In particular, she highlighted how cases like *DeGraffenreid v. General Motors*,[72] doctrinally entrench the existing "political and theoretical approach to discrimination which operates to marginalize Black women" by describing intersectional claimants as seeking a "super-remedy."[73]

D. *Many Scholars Argue That Lower Courts Routinely Misapply the Severe or Pervasive Standard on Summary Judgment*

Several scholars have critiqued the Supreme Court's severe or pervasive standard in Title VII harassment claims, particularly how it has been applied

---

[65] *Id.*

[66] David Benjamin Oppenheimer, *Verdicts Matter: An Empirical Study of California Employment Discrimination and Wrongful Discharge Jury Verdicts Reveals Low Success Rates for Women and Minorities*, 37 U.C. DAVIS L. REV. 511, 536, 539 (2003).

[67] *Id.* at 543.

[68] *Id.* at 544.

[69] *Id.* at 514–15, 552; *see also* Erik Girvan & Heather J. Marek, *Psychological and Structural Bias in Civil Jury Awards*, 8 J. AGGRESSION, CONFLICT & PEACE RSCH. 247, 253 (2016) (finding Black plaintiffs were awarded lower pain and suffering damages than White plaintiffs); Maytal Gilboa, *The Color of Pain: Racial Bias in Pain and Suffering Damages*, 56 GA. L. REV. 651, 655–57 (2022) (summarizing the literature on disparities in pain and suffering awards to Black plaintiffs).

[70] Kimberle Crenshaw, *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*, 1989 U. CHI. LEGAL F. 139, 139–40 (1989) [hereinafter *Demarginalizing the Intersection of Race and Sex*]; Kimberle Crenshaw, *Race, Gender, and Sexual Harassment*, 65 S. CAL. L. REV. 1467, 1470 (1992) [hereinafter *Race, Gender, and Sexual Harassment*].

[71] *Race, Gender, and Sexual Harassment*, *supra* note 70, at 1470.

[72] DeGraffenreid v. Gen. Motors Assembly Div., 558 F.2d 480, 483 (8th Cir. 1977) (quoting lower court opinion characterizing claim brought by Black woman as an attempt to "create a new 'super-remedy'").

[73] Patrick Berning-O'Neill, Note, *"A Reasonably Comparable Evil": Expanding Intersectional Claims Under Title VII Using Existing Precedent*, 24 U. PA. J. CONST. L. 907, 911–12 (2022) (internal quotation marks omitted) (quoting *Demarginalizing the Intersection of Race and Sex*, *supra* note 70, at 141).

by lower courts on summary judgment.[74] Theresa Beiner makes the most sustained critique of the severe or pervasive standard, arguing that the standard "is necessarily a fact-specific inquiry based on social norms that are best assessed by a jury—not one judge sitting in isolation."[75] Beiner catalogued a variety of harassment cases involving strong evidence in the plaintiff's favor, where the court nevertheless granted summary judgment for the defendant.[76] She further argued that lower courts apply the severe or pervasive standard inconsistently or in an overly stringent manner.[77]

Beiner and others have also argued that courts tend to assess each instance of harassing conduct in isolation rather than cumulatively, thereby discounting their severity and frequency. Courts engaging in this piecemeal or "divide and conquer" approach fail to apply the Supreme Court's guidance in *Harris v. Forklift* to consider "all the circumstances."[78] Kenneth Davis catalogued a variety of serious harassment cases dismissed on summary judgment for failing to meet the severe or pervasive standard and argued that the standard "imposes a burden not placed on disparate treatment victims who seek redress under theories other than hostile work environment law."[79] Kleinschmidt argued that several Circuit Courts effectively applied a "severe *and* pervasive" standard.[80] Judith Johnson and others have also argued that courts tend to require more "egregious" facts

---

[74] Judith J. Johnson, *License to Harass Women: Requiring Hostile Environment Sexual Harassment to Be "Severe or Pervasive" Discriminates Among "Terms and Conditions" of Employment*, 62 MD. L. REV. 85, 87 (2003) (arguing that "[i]n no part of Title VII, and for no other term or condition of employment, is the discriminatory conduct required to be severe or pervasive to be actionable").

[75] *Misuse of Summary Judgment*, *supra* note 62, at 75.

[76] *Id.* at 102–19.

[77] Beiner, *supra* note 8, at 809 ("One of the most common rhetorical strategies lower courts use to dispose of sexual harassment cases is simply to declare that no reasonable person could find the behavior sufficiently severe or pervasive.").

[78] *Id.* at 808 (asserting that lower courts "also engage in what I call the 'divide and conquer' approach, whereby rather than looking at the effects of all incidents as the 'totality of the circumstances' standard requires, some courts view the incidents in a piecemeal manner, essentially concluding that each individual instance is insufficient, while failing to consider the cumulative effect of all the incidents"); *see also* Johnson, *supra* note 74, at 123, 129, 131–32 (arguing "[c]ourts [p]arse the [e]vidence to [a]void a [f]inding of [s]evere or [p]ervasive [s]exual [h]arassment" by refusing to consider harassment after the employer took remedial action, or that occurred prior to the 180- or 300-day administrative filing period). Johnson also argues that courts tend to treat evidence of retaliation for complaining as separate from the hostile work environment. *Id.* at 131; Kenneth R. Davis, *The "Severe and Pervers-ive" Standard of Hostile Work Environment Law: Behold the Motivating Factor Test*, 72 RUTGERS U. L. REV. 401, 428 (2020) (stating that some courts "evaluate instances of harassment separately"); *see also* Zimmer, *supra* note 7, at 577 (making a similar argument in the context of discrimination claims, noting "the common practice of courts in slicing and dicing the evidence supporting plaintiff's case in order to grant motions for summary judgment and judgment as a matter of law").

[79] Davis, *supra* note 78, at 426; *see also* Beiner, *supra* note 8, at 809–17 (discussing a variety of cases deemed insufficiently severe or pervasive).

[80] Heather L. Kleinschmidt, *Reconsidering Severe or Pervasive: Aligning the Standard in Sexual Harassment and Racial Harassment Causes of Action*, 80 IND. L.J. 1119, 1130 (2005).

to meet the severe or pervasive standard in sexual harassment cases than in racial harassment cases.[81]

Several commentators have also argued that the objective component of the severe or pervasive test—that a "reasonable person" in the plaintiff's position would need to find the conduct sufficiently severe or pervasive to render the work environment hostile or abusive—should be replaced by a "reasonable woman" standard.[82] Citing social science research suggesting that women are more likely to identify conduct as "harassment," they argue that the legal standard should "capture behaviors that women believe to be sexual harassment."[83] This argument implicitly assumes that *Oncale*'s refinement of the reasonable person standard to include the qualifier "in the plaintiff's position" is insufficient to address differences in experience arising from the plaintiff's identity.[84]

Beiner, Sperino, and Thomas have separately argued that bad precedent imposing an unreasonably high standard for harassment plaintiffs has a

---

[81] Johnson, *supra* note 74, at 88–89, 119–22; Robert J. Gregory, *You Can Call Me a "Bitch" Just Don't Use the "N-Word": Some Thoughts on* Galloway v. General Motors Service Parts Operations *and* Rogers v. Western-Southern Life Insurance Co., 46 DEPAUL L. REV. 741, 742 (1997) (arguing that, in the context of sexual harassment, "the judicial reaction seems less solicitous"); Kleinschmidt, *supra* note 80, at 1125 (making a similar argument, though noting a trend towards demanding more severe conduct in racial harassment cases).

[82] V. Blair Druhan, *Severe or Pervasive: An Analysis of Who, What, and Where Matters When Determining Sexual Harassment*, 66 VAND. L. REV. 355, 365–69, 374 (2013) (discussing the "reasonable woman" standard); Angela Onwuachi-Willig, *What About #UsToo?: The Invisibility of Race in the #MeToo Movement*, 128 YALE L. J. F. 105, 109–10 (arguing that "courts should employ a standard based on a reasonable person in the complainant's intersectional and multidimensional shoes, rather than the ostensibly objective reasonable person standard—which some courts have declared to be male biased") (emphasis omitted). See also Juliano and Schwab's study of cases in the 1990s which found that less than "one quarter of the district court cases mentioned any 'reasonable standard' at all." Juliano & Schwab, *supra* note 56, at 584.

[83] Druhan, *supra* note 82, at 375; *see also* Maria Rotundo, Dung-Hanh Nguyen & Paul R. Sackett, *A Meta-Analytic Review of Gender Differences in Perceptions of Sexual Harassment*, 86 J. APPLIED PSYCH. 914, 919 (2001) (meta-analysis finding that "women are more likely than men to define a broader range of behaviors as harassing"); Kerns, *supra* note 62, at 197 (arguing that the "reasonable woman" standard should apply); Jasmine Tata, *The Structure and Phenomenon of Sexual Harassment: Impact of Category of Sexually Harassing Behavior, Gender, and Hierarchical Level*, 23 J. APPLIED SOC. PSYCH. 199, 205 (1993) (experimental study finding women more likely to identify behaviors as sexually harassing). *Cf.* Elizabeth L. Shoenfelt, Allison E. Maue & JoAnn Nelson, *Reasonable Person Versus Reasonable Woman: Does It Matter?*, 10 AM. U. J. GENDER SOC. POL'Y & L. 633, 666–67 (2002) (noting gender differences in whether respondents believed conduct was harassing, but using the "reasonable woman" rather than the "reasonable person" standard did not produce a clear pro-plaintiff outcome).

[84] Some commentators have concluded that the reference to "plaintiff's position" in the *Oncale* standard implicitly includes the plaintiff's identity. Danielle A. Bernstein, *Reasonableness in Hostile Work Environment Cases After #MeToo*, 28 MICH. J. GENDER & L. 119, 127 (2021) (arguing that *Oncale*'s reference to "in plaintiff's position" served as a "modification" of the "reasonable person" standard, but appeared only as dicta, concluding that "there is still substantial debate over whether the standard should consist of a reasonable person, reasonable woman, or some other version of reasonableness that more explicitly takes into account the victim's characteristics"); ROBERTO L. CORRADA, MICHAEL SELMI, NICOLE BUONOCORE PORTER & MARCIA L. MCCORMICK, EMPLOYMENT DISCRIMINATION LAW 278–79 (11th ed. 2025) (observing that "[t]he Supreme Court ostensibly put this issue [relating to whether a 'reasonable woman' standard should apply] to rest in *Oncale*" by including a reference to the "plaintiff's position"); CATHARINE MACKINNON, SEX EQUALITY 959–60 (2001) (interrogating the "reasonableness" standard).

lasting impact, as courts tend to rely on those rulings years or even decades later.[85] As Beiner argues, "[b]ad [p]recedent [l]eads [t]o [b]ad [p]recedent."[86] Similar to Thomas's argument regarding summary judgment motions generally, Beiner argues that courts should not displace juries on these important questions of fact: "to the extent that harassment law is based on community standards of appropriate behavior in the workplace, by taking these cases away from the jury, judges are not allowing a community standard to develop."[87]

Social scientists have also conducted relevant harassment-related research, principally consisting of studies on how laypeople operationalize "harassment" (typically defined in the study as "sexual harassment").[88] Studies suggest that laypeople consider "physical touching of a sexual nature," and "behaviors that are physically threatening, or involve threats or coercion" to be sexual harassment.[89] As of 2002, the social science literature suggested that laypeople disagreed regarding whether verbal conduct such as "sexual remarks, gestures, sexist jokes, requests for dates, [and] gender harassment" qualified as sexual harassment.[90] However, over half of those surveyed considered these more "ambiguous" or "less obvious" behaviors to be harassment.[91] Laypeople also considered repeated conduct more severe, as well as conduct "targeted" at a specific individual.[92] Drawing on a longstanding survey of federal government workers, Blair Druhan found that respondents were more likely to characterize certain behaviors as harassment if it was committed by a supervisor rather than a co-worker.[93] A meta-analysis found that gender differences in assessments of harassing behavior were larger for "the less extreme and more ambiguous behaviors like derogatory attitudes and dating pressure."[94]

---

[85] SPERINO & THOMAS, *supra* note 50, at 37; Beiner, *supra* note 8, at 817.

[86] Beiner, *supra* note 8, at 817.

[87] *Id.* at 820.

[88] Druhan, *supra* note 82, at 358 (discussing a survey of federal workers regarding "sexual harassment"); Beiner, *supra* note 8, at 821–23 (observing that "[a] great deal of social science research has focused on what sorts of behaviors are deemed sexually harassing," although "[f]ew studies use the legal standard for assessing whether the scenario or particular behavior involved constitutes sexual harassment").

[89] Beiner, *supra* note 8, at 833–34 (summarizing the social science literature).

[90] *Id.* at 834 (citing LOUISE F. FITZGERALD & ALAYNE J. ORMEROD, BREAKING THE SILENCE: THE SEXUAL HARASSMENT OF WOMEN IN ACADEMIA AND THE WORKPLACE 560 (Florence L. Denmark & Michele A. Paludi eds., 1993) (finding that participants were less certain that situations involving gender harassment and less explicit undertones qualified as sexual harassment) ("[C]ontextual variables such as the status of the initiator, the degree of prior relationships, and any so-called suggestive behavior on the part of the woman come into play, influencing the interpretation of these more ambiguous conditions of work. In our perceptions study, we found that the mean ratings of gender harassment items fell into the 'uncertain' range of the scale.")).

[91] Beiner, *supra* note 8, at 834 (internal quotation marks omitted).

[92] *Id.* at 837 (internal quotation marks omitted).

[93] Druhan, *supra*, note 82, at 380–81.

[94] Rotundo, Nguyen & Sackett, *supra* note 83, at 919; Frank E. Saal, Catherine B. Johnson & Nancy Weber, *Friendly or Sexy? It May Depend on Whom You Ask*, 13 PSYCH. WOMEN Q. 263, 272 (1989) (An

II. METHODOLOGY

A. *Overview and Research Questions*

This study contributes to these ongoing scholarly debates in several respects. First, we empirically investigate the following question: do simulated juries reach different conclusions than judges regarding whether the alleged harassment was sufficiently severe or pervasive to qualify as harassment? If study participants conclude that the severe or pervasive standard was met in cases where the court granted summary judgment, then this suggests the courts were misapplying the summary judgment standard or failed to accurately gauge how a reasonable jury would assess the evidence.

Our data also enables us to answer several related questions regarding how judges and simulated jurors assess harassment fact patterns. In particular, we were able to assess:

> (1) Whether courts have become more (or less) stringent in applying the severe or pervasive standard over time, (2) Whether some cases fare better (or worse) than others on the basis of the protected category at issue in the case, and (3) Whether demographic differences among respondents influence their assessments of the severity or pervasiveness of a case.

Lastly, we collected qualitative data by asking the mock jurors to discuss their reasoning regarding their severe or pervasive ratings and ultimate decisions for the case. This methodology provides a more textured basis for understanding any quantitative difference that may be revealed between judges and our simulated juries. Our qualitative analysis can expand our understanding of how laypeople—potential jurors—understand and assess harassing conduct.

B. *Case Sample*

We selected the sample of cases in 2016 by running a Westlaw search for all federal cases using the West key number that identified the severe or pervasive element of a Title VII harassment claim.[95] The search produced 2,843 results. A subset was produced by selecting the first of every 20 cases

---

experimental study by Saal et al. suggests that men are more likely to interpret "a woman's friendly, outgoing behavior as a sign of 'sexiness'" or flirtation).

[95] We used key number 78k1147, which corresponds to the following category: Civil Rights/Employment practices/Harassment; Work environment; Hostile Environment; severity, pervasiveness, and frequency. *West Key Numbering System,* WESTLAW PRECISION, https://1.next.westlaw.com/Browse/SearchKeyNumbers?query=78k1147&jurisdiction=ALLCASES&transitionType=CustomDigest&contextData=(sc.Default) (last visited Sept. 24, 2025). The search produced 2,843 results.

in the search results.[96] The first 100 cases from that subset were then reviewed individually by law student research assistants and coded to identify various features of the case, such as the year it was decided, the procedural posture, the protected categor(ies) at issue, and whether the court deemed the conduct severe or pervasive. Where the case was decided on appeal, we used the appellate ruling on the severe or pervasive element as the case outcome.

We then selected cases for inclusion in the study based on whether the severe or pervasive element was a determining factor in the court's analysis. Cases were excluded from consideration if they were not Title VII harassment cases;[97] were decided on bases unrelated to the severe or pervasive element;[98] or where the court did not specifically discuss the question of whether the conduct was severe or pervasive. Cases involving multiple plaintiffs were also excluded,[99] as were three cases with excessively long fact patterns that could not readily be condensed and two cases where the court did not recount any facts related to the harassment claim.

The initial sample was supplemented in 2020 with additional cases to ensure that more recent cases would be included and that the sample would have sufficient power to assess case year and protected category. This sample was collected using the same Westlaw key number, with separate searches for the 1991–2001 date range, and 2011–2020 range, where 1 of every 10 cases was selected for review.[100] An additional subset of 28 cases was collected and assessed for inclusion using the criteria described above.

[96] For example, result number 1, 21, 41, 61, 81, etc.

[97] Several cases classified under the applicable Westlaw key note were retaliatory harassment cases, where the hostile work environment was not based on a protected category but in response to protected activity.

[98] Some cases were dismissed on multiple bases but were included if there was a discussion of the severe or pervasive element.

[99] Cases with multiple plaintiffs were not suitable for inclusion because fact patterns often involved intertwined allegations that could not be readily condensed for individual plaintiffs.

[100] Here, result number 3, 13, 23, etc. were selected for download, with 12 cases downloaded from the 1991–2001 date range, and 16 cases downloaded from the 2011–2020 date range.

Table 1. Summary of Case Sample.

Decade	**Number of cases**	**Proportion of total**
1995-1999	6	0.07
2000-2009	42	0.52
2010-2019	33	0.41
Protected category		
Race	25	0.31
Sex	27	0.33
Other (e.g. religion, color, national origin)	5	0.06
Intersectional	24	0.30
Employer Type		
Public sector	36	0.44
Private sector	45	0.56
Procedural Posture		
Motion for summary judgment (or related appeal)	75	0.91
Motion to dismiss	5	0.06
Motion for judgment as a matter of law (appeal)	1	0.01
Jury verdict (appeal)	1	0.01
Total observations	81	

The final case sample ultimately included 81 unique cases, decided between 1995 and 2019.[101] Of these, 76 involved a motion for summary judgment—or an appeal of a motion for summary judgment—and 5 involved a motion to dismiss. All cases in the sample included some determination on whether there was an issue of fact on the question of whether the conduct qualified as severe or pervasive.[102] We aggregated both types of motions together, as courts appeared to approach them in a broadly similar manner notwithstanding the procedural differences between the two types of motions.[103] Of these, the courts declared 28 out of 81 (35%) sufficiently severe or pervasive to create an issue of fact, allowing those cases to proceed. In the remaining 53 out of 81 cases (65%), the court concluded that no reasonable juror would find that the conduct was severe or pervasive and granted summary judgment or a motion to dismiss for the defendant. This proportion is similar to the rate observed in broader samples of employment discrimination cases.[104]

C. *Participant Sample*

Participants were recruited through Amazon's Mechanical Turk (MTurk) platform in 2019, 2020, and 2022, and paid approximately $1.25 each.[105] Following initial rounds of survey data collection in 2019 and 2020, subsequent data collection in 2020 and 2022 was limited to race-based, sex-based, and intersectional claims based on multiple protected characteristics, to yield enough statistical power to analyze how each protected category may influence case outcomes. Within those categories, intersectional claims

---

[101] One of the 81 cases was not randomly assigned to respondents by the Qualtrics algorithm and received zero ratings. Consequently, virtually all of the quantitative analysis presented herein refers to a sample of 80 cases.

[102] There were a few cases where the court found the conduct sufficiently severe or pervasive to create an issue of material fact, but granted summary judgment on other grounds, such as the absence of vicarious liability. *See, e.g.*, Henneman v. Airtran Airways, 705 F. Supp. 2d 1012, 1034–35 (E.D. Wis. 2010). In such instances, the case was coded as surviving summary judgment as to the severe or pervasive element.

[103] For the few cases involving a motion to dismiss, the Courts assessed whether the facts as alleged in the Complaint would be sufficient to qualify as "severe or pervasive." *See, e.g.*, Edwards v. Murphy-Brown, LLC, 760 F. Supp. 2d 607, 630 (E.D. Va. 2011); Prince v. Madison Square Garden, 427 F. Supp. 2d 372, 377 (S.D.N.Y. 2006).

[104] *See* Berger, Finkelstein & Chung, *supra* note 53, at 54–55, 60 (study finding 36.4% of employment discrimination cases—including harassment cases—on PACER survived summary judgment, while 30.9% of equivalent published cases survived summary judgment. The sample included 21 sex-based harassment cases, which survived summary judgment 52% of the time.); Weinberg & Neilsen, *supra* note 54, at 332, 336 (reviewing summary judgment rulings in employment civil rights cases; court granted full summary judgment in 56% of cases, while at least some portion of the case survived in 44% of cases). Our study focused only on whether the harassment claim survived summary judgment—and not whether other claims survived. Thus, some of the cases coded as losing on summary judgment may have survived as to claims other than harassment and would have been coded as surviving summary judgment in the Weinberg and Nielsen study. *See also* Clermont & Schwab, *supra* note 61, at 440 (finding that approximately 70% of employment discrimination and harassment cases settled).

[105] Approval was obtained from the institutional review board at the University of Oregon and Georgetown University.

were oversampled, as they appeared multiple times in the survey—once for each protected category at issue.[106]

Table 2, below, compares the demographics of the sample to the U.S. population. Women were underrepresented, making up 36% of the sample compared to 51% of the U.S. population.[107] College graduates were overrepresented in the sample (49% versus 34% of the U.S. population), and high school graduates underrepresented (8% versus 37% of the U.S. population).[108] The age demographics of the sample were also younger than the U.S. population—77% of the sample was between 18 and 44 years old, compared to 37% of the U.S. population, presumably due to the online platform through which responses were gathered.[109] The race-based demographics were comparable to the U.S. population, although Latino respondents were somewhat underrepresented (8% of the sample compared to 19% of the population). The political views of the sample population were comparable to the United States as a whole, with the sample roughly equally divided.

[106] Some intersectional claims were based on race or sex and another protected category. These claims were included in the 2020 and 2022 data collection.

[107] *Age and Sex Composition: 2020*, U.S. CENSUS BUREAU, https://www.census.gov/library/publications/2023/decennial/c2020br-06.html (Dec. 10, 2025).

[108] *Educational Attainment in the United States: 2020*, U.S. CENSUS BUREAU, https://www.census.gov/data/tables/2020/demo/educational-attainment/cps-detailed-tables.html (Dec. 10, 2025) (Table 2. Education Attainment of the Population 25 Years and Over, by Selected Characteristics: 2020, Both Sexes).

[109] *Age and Sex Composition: 2020,* U.S. CENSUS BUREAU, *supra* note 107.

Table 2. Descriptive Statistics of Sample vs. U.S. Population

Characteristic/Variable	**Sample**[110]	**U.S. Population**[111]
Female	0.36	0.51
Educational attainment		
High school graduate or less	0.08	0.37
Some college	0.19	0.15
College graduate	0.49	0.34
Graduate degree	0.22	0.14
Age bracket[112]		
18–44	0.77	0.37
45–64	0.20	0.26
65+	0.02	0.13
Race[113]		
White	0.66	0.58
Latino or Hispanic	0.07	0.19
Black or African American	0.15	0.12
Asian	0.08	0.06
Native American/Pacific Islander	0.01	0.01
Multiracial	0.04	0.10
Political views		
Liberal	0.33	0.25
Moderate[114]	0.30	0.37
Conservative	0.35	0.36
Total observations	699	

[110] Figures add up to less than 100% for the sample where not every respondent answered all demographic questions.

[111] *Race and Ethnicity*, U.S. CENSUS BUREAU, https://data.census.gov/profile/United_States?g=010XX00US#race-and-ethnicity (last visited Apr. 1, 2024).

[112] U.S. CENSUS BUREAU, *supra* note 107.

[113] The United States Census figure adds up to more than 100% because the "Latino or Hispanic" category includes respondents of any race who also identify as Latino or Hispanic (for example, Black and Hispanic). For the purposes of Table 2, the "White" category excludes those who identify as Latino or Hispanic. Those who identify as both White and Latino/Hispanic comprise about 4% of the United States population. The Census also includes an additional, "Some other race" category, which is not listed in Table 2 and comprises 8% of the United States population.

[114] From our survey, the "Moderate" category includes those who responded that they were "slightly liberal" and "slightly conservative" as well as thought who identify as "moderate; middle of the road" and "haven't thought much about this."

D. *Research Designs and Procedures*

We used an online survey using Qualtrics software to gather participants' assessments of the selected cases. The first step in our study design was to select the fact patterns from the cases that would be reviewed and rated by the participants. We reviewed the full judicial opinions for the 81 cases and excerpted the fact section to a length that was approximately 500 words or less. We did this to increase the readability of the facts for respondents in the context of our online survey. None of the language in the case relating to the law was included in the excerpt, only the facts related to the alleged harassment.

The Qualtrics software randomly assigned each of the 699 participants to review one single fact pattern from the case sample. Once participants opened the survey link, they consented to participate in the study and then were presented with one of the excerpted fact patterns. The fact pattern was prefaced with the following introductory sentence, with the relevant bracket information filled in:

> The plaintiff is [name]. The defendant is [name of employer]. Plaintiff alleges harassment on the basis of [protected category].[115]

Participants were also given excerpts of the jury instructions for Title VII harassment claims.[116] Participants were only presented with one single protected category in the introductory sentence. Intersectional cases thus appeared multiple times in the sample, referring only to one protected category at a time. So, for example, in a case that alleged both sex and race discrimination, one respondent might be told the case involved race-based harassment, while another respondent might be instructed that the case involved sex-based harassment. This approach mirrors the practical reality that Title VII plaintiffs are expected to plead separate claims for each protected category at issue.

The survey then asked participants to: (1) rate how severe or pervasive the fact pattern was on a scale of 0–100; (2) assess whether the severe or pervasive standard was met based on the jury instructions provided ("Yes" or "No");

---

[115] For example: "The plaintiff is **Rasmy**. The defendant is **Marriott**. Plaintiff alleges harassment on the basis of **religion**."

[116] The excerpt read as follows (tailored in this case for race-based harassment): "Your verdict must be for the plaintiff and against the defendant on the plaintiff's claim of race-based harassment if all of the following elements have been proved:

First, the plaintiff was subjected to workplace conduct or conditions alleged by the plaintiff; and
Second, such conduct was unwelcome; and
Third, such conduct was based on the plaintiff's race; and
***Fourth*, such conduct was sufficiently severe or pervasive that a reasonable person in the plaintiff's position would find the plaintiff's work environment to be hostile or abusive; and**
Fifth, at the time such conduct occurred and as a result of such conduct, the plaintiff believed his/her work environment to be hostile or abusive.

If any of the above elements has not been proved, your verdict must be for the defendant and you need not proceed further in considering this claim."

and (3) provide a narrative answer to explain their responses. The final section of the survey included a demographic questionnaire. Upon completion of the study, participants were given an attention check question asking the name of the plaintiff in the case. Responses were initially reviewed for completeness, producing an initial sample of 928 responses.[117] They were then individually reviewed to assess whether the responses to the attention check and narrative question were completed by a human being who understood the question being asked, and not a robot. Nonsensical or non-responsive submissions were removed, producing a final sample of 699 responses.

The combined 699 participant surveys and sample of 81 Title VII harassment cases produced a nested data structure, with multiple ratings for each case in the sample. On average, each case was rated 8.5 times, although the number of ratings varied due to the random case assignment by the Qualtrics software.[118]

E. *Measurement*

An empirical answer to the question of what a "reasonable jury" may decide first involves the threshold question of the level of agreement—or disagreement—among simulated jurors that is equivalent to a reasonable jury.[119] Civil procedure scholars have not answered this question with any precision because summary judgment itself involves only a single decisionmaker (the judge) attempting to guess how a reasonable jury might assess the evidence.[120]

One potential option in simulating how a reasonable jury might decide a case would be to apply the legal rule for the level of consensus required of a jury in that particular case. In Title VII cases, that standard is unanimity.[121] However, the unanimity standard applies post-deliberation—juries are only expected to reach a unanimous verdict following extensive deliberation, where disagreement is discussed and resolved over time. It does not reflect

---

[117] It is common for surveys on MTurk to be abandoned after partial completion.

[118] As previously noted, one of the cases received no ratings due to the random case assignment by the Qualtrics software. In addition, intersectional cases received more ratings on average because they appeared multiple times in the survey (once for each protected category at issue).

[119] Suja Thomas argues that both the Supreme Court and lower courts have been imprecise even in the standard to apply, using the terms "reasonable jury," "reasonable juror," "rational juror," and "rational factfinder" somewhat "interchangeabl[y]" even though the different formulations would likely produce different outcomes. Thomas, *supra* note 22, at 776. Thomas ultimately recommends that courts apply a "reasonable jury" standard, although she does not operationalize how that standard might work. *Id.*

[120] A few commentators offer relevant commentary without answering the more quantitative question. Theresa Beiner's assessment of summary judgment cases in the harassment context suggests perhaps that summary judgment would rarely—or perhaps never—be appropriate on the basis of the "severe or pervasive" element because the question "is necessarily a fact-specific inquiry based on social norms." *Misuse of Summary Judgment*, *supra* note 62, at 75. Kahan makes the more general recommendation that courts adopt a stance of "judicial humility" where they perform a "mental double check when ruling on a motion that would result in summary adjudication" and question whether certain subcommunities would reach a different conclusion on the same set of facts. Kahan, Hoffman & Braman, *supra* note 37, at 897–99.

[121] FED. R. CIV. P. 48(b).

a sample like ours, where respondents provide a rating without the opportunity to deliberate. A unanimity-based standard is also poorly suited to identify fact patterns where jurors exhibit a high level of disagreement about whether the standard is met. The presence of such disagreement suggests there is a disputed issue of fact for the jury to consider, and the case should not be dismissed on summary judgment.

A second option is to set the reasonable jury threshold at the lowest possible level, where at least one respondent in the study thought the severe or pervasive standard was met in a particular case. However, a *de minimis* threshold is potentially inconsistent with the legal standard for summary judgment. The "reasonable" qualifier in the reasonable jury standard assumes that an individual juror (or even group of jurors) might not be reasonable. If a lone respondent believes the standard was met, while 80% or 90% of others disagree, it suggests both that the individual's assessment is idiosyncratic and that it is unlikely that a collective reasonable jury would reach the same conclusion as that lone juror following deliberation. The Supreme Court's decision in *Scott v. Harris* also contemplates that extremely lopsided evidence in the defendant's favor compels summary judgment.[122]

Taking all these factors into account, we provide two possible benchmarks to approximate a reasonable jury verdict in the plaintiff's favor: (1) where at least a substantial minority (40%) of raters believe the severe or pervasive threshold has been met, and (2) where at least half (50%) of raters believe the threshold has been met. Where a critical mass of raters concludes the standard has been met, those simulated jurors would stand a fair chance of persuading other jurors that the plaintiff has met their ultimate burden of proof upon deliberation. Thus, we reason that with both of these thresholds, a reasonable jury may indeed find in favor of the plaintiff.

III. QUANTITATIVE ANALYSIS AND RESULTS

A. *Are Courts Dismissing Cases on Summary Judgment That a Reasonable Jury Would Consider Severe or Pervasive?*

There was a high degree of variance in respondent ratings, both within and between cases. On average, respondents gave the cases they reviewed a mean rating of 58 out of 100 on the "severe or pervasive" question, and a median rating of 64. Even ratings within cases had a high degree of variance, the average standard deviation within cases was 22.7 units on the severe or pervasive rating scale. A high level of variance within cases suggests that respondents may be interpreting the same set of facts differently, where some respondents consider the facts more serious than others. Within-case variance

[122] *See* Scott v. Harris, 550 U.S. 372, 380–81 (2007) ("Respondent's version of events is so utterly discredited by the record that no reasonable jury could have believed him.").

may also be a function of the 0–100 scale—a wide range that gave respondents considerable discretion within the parameters of the jury instructions.

Figure 1 is a boxplot illustrating the range of case ratings, sorted by the outcome of the case. The plot on the left includes participant ratings of cases where the court granted summary judgment for the employer. The plot on the right includes the ratings of cases where the court ultimately denied summary judgment and permitted the case to proceed. Each dot represents an individual rating of a single case by a respondent. The Y-axis consists of the respondent's severe or pervasive rating on the 0–100 scale, where a 0 is "not at all" and 100 is "extremely" severe or pervasive harassment.

Most of the data appears on the left plot, as the court granted summary judgment in most of the cases included in the study. On average, the cases that survived summary judgment were rated as 71 out of 100 on severity. The cases that did not survive summary judgment had an average severity rating of 53 out of 100.

The density of data on the top left quadrant of the boxplot (ratings above 76 for cases dismissed on summary judgment) suggests that respondents considered many cases that were dismissed to be quite severe or pervasive. There were also a fair number of additional dismissed cases with reasonably strong ratings: 25% of the ratings for dismissed cases were rated between 59.5 and 76 out of 100.

At the same time, the light density in the bottom right quadrant of Figure 1 suggests that courts and respondents largely agreed regarding cases that survived summary judgment. Only 25% of the ratings for cases that survived summary judgment were below 60.5.

Figure 1. Severe or pervasive ratings by case outcome.

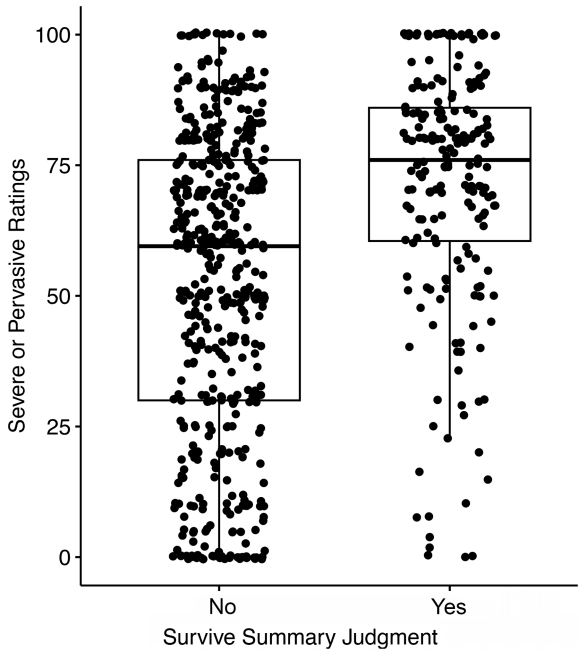

Table 3 displays the proportion of cases that would have survived summary judgment based on the thresholds identified in Part II.E, above—that is, the number of cases where 40% or 50% of jurors concluded that the severe or pervasive standard was met. The vast majority of cases would have survived summary judgment based on this standard. In 67 out of the 75 (89%) cases with at least 4 ratings, at least 40% of the raters thought the severe or pervasive threshold was met. Similarly, in 62 out of 75 cases (83%), at least half of the respondents thought the severe or pervasive standard was met. By comparison, only 35% of the cases in the sample actually survived a motion for summary judgment or motion to dismiss. These results suggest that courts are far too aggressive in dismissing cases on the basis of the severe or pervasive element, and that in many such cases, a reasonable jury very well may have found the facts sufficiently severe or pervasive had the case proceeded to trial.

Conversely, there were a small number of cases where most jurors believed that the severe or pervasive standard was not met, and for which summary judgment was proper. The challenge for judges—which we address in greater detail in Parts IV and V below—will be to distinguish between the vast majority of cases unsuitable for summary judgment on the basis of the severe or pervasive threshold, and the rare cases that a jury would consider frivolous.

Table 3. Percentage of cases that would have survived dismissal based on mock juror assessments.

	YES (Survive motion)[123]	NO (Dismissal)
Court thought the severe or pervasive standard was met in the case	26 cases (35%)	49 cases (65%)
> 50% of respondents thought the severe or pervasive standard was met in the case	62 cases (83%)	13 cases (17%)
> 40% of respondents thought the severe or pervasive standard was met in the case	67 cases (89%)	8 cases (11%)

N cases = 75 [124]

B. *Have Courts Become More (or Less) Stringent in Applying the Severe or Pervasive Standard Over Time?*

The nature of this dataset enabled us to use regression analysis to assess whether courts have altered their threshold for what qualifies as severe or pervasive harassment over time. The independent variables of interest include the date the case was decided and the mean respondent ratings for the severity or pervasiveness of the conduct. We used a logistic regression to test whether courts were granting summary judgment more or less over time, controlling for the independently collected severe or pervasive ratings. It is possible, for example, that courts have become more rigid over time, allowing fewer cases to survive summary judgment than in earlier decades, after controlling for severity.

Our logistic regression results reveal that the year of the case was not a significant predictor of case outcome, after controlling for the mean severe or pervasive ratings (see Table 4). Consequently, the results do not support

[123] Here we define "survive summary judgment" to mean that the court found an issue of fact on the question of whether the severe or pervasive element was met. Our statistics do not take into account whether the court granted summary judgment based on other elements of the claim.

[124] Only includes cases with four or more ratings.

the hypothesis that courts have become more or less likely to grant summary judgment over time.

Notably, the mean severe or pervasive rating was a statistically significant predictor of case outcome (β=0.125, $p<0.01$). A 1-point increase in mean severe or pervasive rating (on a scale of 100) would increase the odds of surviving summary judgment by 13.3%. The model in Table 4 produced an adjusted R squared of 0.329, indicating that 33% of the variance in case outcomes can be predicted by the severe or pervasive ratings. In other words, the severe or pervasive rating was a strong and significant predictor of whether a case survives on summary judgment.

Table 4. Logistic Regression Analysis of Case Outcomes[125]

	Dependent variable:
	Case outcome (1= survive summary judgment)
Mean severe or pervasive rating	0.125***
	(0.029)
Year	0.064
	(0.053)
Constant	-137.539
	(106.785)
Observations	80
Adjusted (or Pseudo) R^2	0.329

* p<0.1; ** p<0.05; *** p<0.01

We also conducted an ordinary least squares regression to analyze whether the year of the case predicted the mean severe or pervasive ratings. This was to test whether the fact patterns describing the alleged harassment have become more or less severe or pervasive over time, based on the participants' assessments (see Table 5). The "year" predictor was not significant, suggesting that types of alleged harassment that defendants challenge as insufficiently severe or pervasive have remained similar over time.[126]

[125] This regression model consisted of a logistic generalized linear model with binomial family.

[126] We also conducted a multilevel regression to assess whether severe or pervasive ratings were affected by the year the data was collected for the study: 2019, 2020, and 2022, perhaps due to cultural changes in connection with COVID or a continuing effect of the #MeToo or #BlackLivesMatter movements. The collection date did not have a statistically significant effect on case ratings.

Table 5. Linear Regression Analysis of Mean Severe or Pervasive Rating.

	Dependent variable:
	Mean severe or pervasive rating
Year	-0.196
	(0.331)
Constant	454.649
	(664.301)
Observations	80
Adjusted (or Pseudo) R^2	0.004

* p<0.1; ** p<0.05; *** p<0.01

C. *Do Some Cases Fare Better (or Worse) Than Others on the Basis of the Protected Category at Issue?*

Next, we used logistic regression to analyze whether some types of harassment cases fare better than others on summary judgment, holding constant the severe or pervasive ratings. We focused this analysis on the protected category that the alleged harassment was based on, including "Race only," "Sex only," "Other," and "Intersectional." We combined cases involving age, religion, and national origin-based harassment into the "Other" category because the sample of those cases was too small to detect statistically significant differences across those specific types of cases. "Intersectional" cases included any case where the plaintiff alleged the harassment was based on more than one protected category. The "Intersectional" category extends beyond cases alleging both race and sex-based harassment and includes cases alleging other combinations, such as religion and race, or age and sex. In our logistic regression analysis, the case type is coded as a dummy variable, with "sex cases" as the reference category. Thus, a statistically significant finding indicates that the outcome of the case type statistically differs from the outcome for sex cases in the dataset.

The logistic regression results presented in Table 6 reveal that intersectional cases involving more than one protected category are less likely to survive summary judgment than sex-based harassment cases, even after controlling for the mean severe or pervasive ratings for a given case (Odds ratio=0.146, β=-1.920, $p<0.05$). The effect size for intersectional cases is quite large; even after controlling for the mean severe or pervasive ratings (Model 2), intersectional cases have about 85.4% lower odds of surviving summary judgment than similarly severe sex-based harassment cases. Race-based cases also fared worse than sex-based cases, but the effect was only marginally significant in Model 1 (Odds ratio=0.387, β=-0.950, $p<0.1$) and was not statistically significant when the mean severe or

pervasive ratings were added in Model 2. Follow-up analysis should be conducted with a larger sample of cases to further explore these differences.

Table 6. Logistic Regression Analysis Predicting Case Outcomes, by Protected Category.

	Dependent variable:	
	Survive Summary Judgment	
	(1)	(2)
Intersectional case (β_1)	-2.773***	-1.920***
	(0.836)	(0.930)
Other case (β_2)	-1.761	0.293
	(1.185)	(1.386)
Race case (β_3)	-0.950*	-0.754
	(0.572)	(0.692)
Mean Severe or Pervasive rating (β_4)		0.117***
		(0.033)
Constant (β_0)	0.375	-7.603***
	(0.392)	(2.248)
Observations	81	80
Pseudo R^2	0.158	0.368

Robust standard errors in parentheses

Reference category is "Sex case"

* p<0.1; ** p<0.05; *** p<0.01

We considered the possibility that the type of employment (public versus private sector) might have served as a confounding variable, as several of the cases with poor ratings from our simulated juries were brought against government employers. However, we included a variable for government employment as a control variable in several regressions, and the results were not statistically significant.

D. *Do Respondent Demographics Predict Their Assessment of the Severity or Pervasiveness of a Case?*

We then fit a linear model with fixed effects, using the case's mean severe or pervasive rating as the outcome variable. The model includes demographic predictors and case-level fixed effects. The demographic predictors included in the model included participants' education, age, gender, political views, and race.

The model formula is given by:

$$Y_{i,c} = \beta_0 + \sum_{j=1}^{2} \beta_{1j}\text{Education}_{i,j} + \sum_{k=1}^{2} \beta_{2k}\text{Age}_{i,k} + \beta_3\text{Female} + \sum_{m=1}^{2} \beta_{4m}\text{PoliticalViews}_{i,m} + \sum_{n=1}^{5} \beta_{5n}\text{Race}_{i,n} + \gamma_c + \varepsilon_{i,c}$$

Where:

- $Y_{i,c}$ is the severity or pervasiveness rating for respondent *i* in case c
- β_0 is the intercept
- γ_c is the fixed effect for each case
- $\varepsilon_{i,c}$ is the error term

Surprisingly, few demographic characteristics had a statistically significant effect on severe or pervasive ratings (see Table 7). Although extant literature suggests that women perceive harassing behaviors differently than men,[127] results suggest that women did not differ significantly in their ratings of the severe- or pervasive-ness of fact patterns. The results also suggest that political views did not have a significant effect on case ratings.

[127] *See* Druhan, *supra* note 82, at 365–69, 374 (describing the "reasonable woman standard" and subsequent results); Onwuachi-Willig, *supra* note 56, at 109–10 (noting that courts should employ a standard based on the complainant's perspective); Juliano & Schwab, *supra* note 82, at 552, 582 (same).

Table 7. Regression Results: OLS With Fixed Effects.

	Dependent variable:
	Severe or Pervasive Rating
Education: College Graduate (β_{11})	4.942**
	(2.454)
Education: Graduate Degree (β_{12})	8.022***
	(3.018)
Age: 45-64 (β_{21})	-0.601
	(2.668)
Age: 65+ (β_{22})	19.137***
	(7.301)
Gender: Female (β_{3})	0.359
	(2.171)
Political Views: Conservative (β_{41})	-1.782
	(2.300)
Political Views: Moderate (β_{42})	-4.668
	(3.371)
Race: Asian (β_{51})	-3.066
	(4.099)
Race: Black or African American (β_{52})	7.906***
	(3.029)
Race: Latino or Hispanic (β_{53})	3.847
	(4.166)
Race: Multiracial (β_{54})	5.139
	(5.362)
Race: Native American or Pacific Islander (β_{55})	-23.882***
	(8.519)
Case Fixed Effect	Yes
Observations	665
R^2	0.352
Adjusted R^2	0.249

Note: *p<0.1; **p<0.05; ***p<0.01

Black respondents rated harassment as more severe than White respondents, on average rating cases 8 points higher on a 100-point scale, a difference that was statistically significant (β=7.906, p<0.01). This finding is consistent with prior research.[128] Respondents that identified as Native American or Pacific Islander tended to give lower ratings than White

[128] *See* Weinberg & Nielsen, *supra* note 54, at 339 (finding that White judges were more likely to grant summary judgment than their minority counterparts).

respondents, and this effect was statistically significant (β=-23.882, $p < 0.01$). However, the sample included only 10 respondents who identified as Native American or Pacific Islander, and their individual assessments may not be representative of how a larger group might have rated the same cases. Other racial categories did not produce significant results compared to White respondents.

With respect to age, respondents aged 45–64 did not statistically differ from those under age 45 in their ratings. However, respondents aged 65 or older rated fact patterns are more severe or pervasive, on average rating cases 19 points higher on a 100-point scale than younger respondents (β=19.137, $p < 0.01$). This finding is somewhat counterintuitive and could reflect the idiosyncratic preferences of the small number of older respondents who participated in the study.

Education was also associated with higher ratings on the severe or pervasive scale. College graduates, on average, rate cases as 5 points more severe or pervasive than those without a college degree (β=4.942, $p < 0.05$). Respondents with a graduate degree rate cases as 8 points more severe or pervasive than their high school counterparts (β=8.022, $p < 0.01$).

IV. QUALITATIVE ANALYSIS AND RESULTS

We also gathered qualitative data to gain insights into the mock jurors' decision-making process. This stage of data collection is designed to further enrich our understanding and interpretation of the quantitative findings. After respondents were asked whether the plaintiff proved that the conduct was sufficiently severe or pervasive to meet the legal standard, we then asked them to "Please explain why." Respondents were provided with a text box to share their reasoning.

Narrative answers were first sorted based on whether the respondent separately answered "yes" or "no" to the question of whether the conduct was sufficiently severe or pervasive. The narrative answers were then reviewed and grouped based on observed themes and trends in their responses.

Below, we report the broad themes that emerged from the narrative responses, and what they suggest about how jurors assess harassment evidence. We also compare and contrast respondents' analysis with the actual court analysis in the case—information that was not known to the respondents when they completed the survey.

Sections A through E, below, describe respondents' reasoning in support of their conclusions that the severe or pervasive standard was met. Section F attempts to identify commonalities among the 8 cases that fewer than 40% of respondents concluded met the severe or pervasive standard, informed by respondents' narrative responses.

Broadly, the results suggest that respondents viewed the facts far more holistically than courts. Psychologist William James described expertise as

"the art of knowing what to overlook"—"a genuine dropping out and throwing overboard of conscious content."[129] He noted that many talented experts "are notoriously bad lecturers" because "they never spontaneously see the subject in the minute articulate way in which the student" does.[130] Similarly, our qualitative results suggest that judges tend to discount swaths of evidence that our simulated jurors considered important in assessing how the plaintiff likely experienced the workplace. Like James's experts, judges "overlook" much of what matters to jurors and consequently fail to accurately predict how a jury would respond to the evidence presented.

A. *The* Harris v. Forklift *Factors*

1. *"All the circumstances"*

Respondents were not provided with the Supreme Court's language in *Harris v. Forklift*, advising lower courts to consider "all the circumstances" including the "frequency of the discriminatory conduct; its severity; whether it is physically threatening or humiliating, or a mere offensive utterance."[131] Nevertheless, many of the respondents considered these factors—perhaps reflecting the lawyers' or court's emphasis on facts that were the most legally relevant.[132] As described in greater detail below, many respondents referenced the frequency of the conduct and considered physical contact or threats to be particularly offensive. They also made distinctions between "mere offensive utterance[s]" and genuinely hostile work environments.

Some respondents identified the frequency or the seriousness of the conduct as the basis for their decision—although others tended to be more focused on the severity of the conduct than its frequency. A few respondents cited the prolonged period over which the conduct occurred (e.g., "He was being discriminated [against] for a six month time period so it was an ongoing issue"[133]). Another referenced the total number of harassing incidents (e.g., "30 times is enough consecutive times to constitute repeated and that there was intended harassment"[134]), while another considered both the time period and number of instances ("It was multiple things that were done to offend him. It was done over and over in a six-month period."[135]).

In describing the severity, many respondents referenced whether the harasser made physical contact with the plaintiff or whether the conduct

---

[129] 2 WILLIAM JAMES, THE PRINCIPLES OF PSYCHOLOGY 369 (Dover Publ'ns 1950) (1890).

[130] *Id.*

[131] Harris v. Forklift Sys., Inc., 510 U.S. 17, 23 (1993).

[132] It is also possible that the *Harris v. Forklift* factors could reflect intuitive notions of what makes a work environment hostile and abusive.

[133] Respondent discussion of the facts in Honor v. Booz-Allen & Hamilton, Inc., 383 F.3d 180, 183–84, 191 (4th Cir. 2004) (affirming summary judgment as to harassment claim).

[134] Respondent discussion of the facts in Goller v. Ohio Dep't of Rehab. & Corr., 285 F. App'x 250, 252, 259 (6th Cir. 2008) (affirming summary judgment as to harassment claim).

[135] Respondent discussion of the facts in Jones v. Delta Towing LLC, 512 F. Supp. 2d 479, 485, 493 (E.D. La. 2007) (denying summary judgment as to harassment claim).

could be interpreted as a threat. Some respondents cited a single act of physical harm or contact as the basis for their determination that the standard was met (e.g., "according to him his left [sic] left eye was harmed"[136]; "She was assaulted while standing in the hallway, minding her own business"[137]; "She was drugged and raped while on a company layover"[138]; "Because he physically grabbed her and caused her stress and hardship"[139]). Other respondents also highlighted the severity of overt or implied threats (e.g., "It happened repeatedly, and at least once he had to fear he was about to be physically assaulted"[140]; "Finding a noose in the workstation, implies to me that the workplace is dangerous and a place where Taylor may not feel safe due to her skin color"[141]).

Some of the corresponding court rulings denied summary judgment on similar grounds to the bases described by respondents.[142] Others, however, paid lip service to the Supreme Court's guidance to consider "all the circumstances," but—to borrow Michael Zimmer's phrase—tended to "slice and dice" the evidence.[143] In *Goller v. Ohio Dep't of Rehab. & Corr.*, for example, the court acknowledged the plaintiff had been called "barbie" thirty times, but concluded the plaintiff had been insufficiently specific about the number of times she had been called names where the "offensive term [was] proceeded by the modifier 'white.'"[144] The court also ignored all other additional instances of harassment alleged by the plaintiff.[145]

Other corresponding court rulings granted summary judgment for the employer by discounting the severity of the conduct, and comparing the facts to other, more egregious cases where summary judgment was granted. *Yang v. Lakewood Management*, for example, involved allegations of physical contact, including one instance where "pepper sauce was put on plaintiff's head" and dripped into his eye, and a separate incident where his "left eye

---

[136] Respondent discussion of the facts in Yang v. Lakewood Mgmt., LLC, 918 F. Supp. 2d 1205, 1206, 1209 (D. Kan. 2013) (granting summary judgment as to harassment claim).

[137] Respondent discussion of the facts in Tolson v. Springer, 618 F. Supp. 2d 14, 19, 24 (D.D.C. 2009) (granting summary judgment as to harassment claim).

[138] Respondent discussion of the facts in Ferris v. Delta Air Lines, Inc., 277 F.3d 128, 131–32, 138 (2d Cir. 2001) (vacating summary judgment as to harassment claim).

[139] Respondent discussion of the facts in Sowemimo v. D.A.O.R. Sec., Inc., 43 F. Supp. 2d 477, 482, 491–92 (S.D.N.Y. 1999) (denying summary judgment as to harassment claim).

[140] Respondent discussion of the facts in Jones v. UPS Ground Freight, 683 F.3d 1283, 1293 n.34, 1304 (11th Cir. 2012) (vacating summary judgment for a harassment claim).

[141] Respondent discussion of the facts in Taylor-Norman v. Joco Assembly, No. 09-CV-410-TCK-FHM, 2010 WL 3521610, at *7 (N.D. Okla. Sept. 7, 2010) (granting motion to dismiss as to harassment claim).

[142] *See, e.g.*, *Sowemimo*, 43 F. Supp. 2d at 492 (denying summary judgment as to harassment claim); Jones v. Delta Towing LLC, 512 F. Supp. 2d 479, 487, 493 (E.D. La. 2007) (same); *Jones*, 683 F.3d at 1298 (same).

[143] Zimmer, *supra* note 7, at 577. As discussed above, Teresa Beiner leveled this critique specifically in the context of harassment cases decided on summary judgment, characterizing the judicial approach as "divide and conquer." Beiner, *supra* note 8, at 808.

[144] Goller v. Ohio Dep't of Rehab. & Corr., 285 F. App'x 250, 259 (6th Cir. 2008).

[145] *Id.*

was harmed by kitchen chemicals."[146] Seventy-five percent of respondents who reviewed the case concluded the plaintiff met the severe or pervasive standard. The district court, however, did not, citing *Faragher v. Boca Raton* for the proposition that "courts should filter out complaints which rely upon isolated incidents, sporadic jokes and occasional teasing."[147] (The district court, however, omitted related language in *Faragher*, clarifying that "extremely serious" isolated incidents could qualify as harassment.)[148] The court also compared the facts in *Yang* to a Tenth Circuit case where summary judgment was granted for the employer despite evidence that a surgeon made demeaning comments toward a nurse, flicked her on the head, and threw a piece of human tissue at her leg.[149]

Similarly, the court granted summary judgment for the employer in *Tolson v. Springer*, which involved quite a long list of racially harassing incidents by a co-worker, including threats and acts of physical aggression. All six of the respondents who reviewed the facts in *Tolson* concluded that it met the severe or pervasive standard. Here too, the court justified its ruling with reference to three other extreme cases from the D.C. Circuit where summary judgment was granted for the employer.

Likewise, in another case—*Taylor-Norman v. JoCo*[150]—all the respondents in our study concluded the conduct was severe or pervasive, and a number highlighted the threatening nature of a noose planted in the workplace. However, the court granted a motion to dismiss the case on the basis that the noose was intended for a different employee and completely overlooked the effect that the noose might have on other Black workers. The only legal precedent the court offered to support its conclusion was language from the Tenth Circuit, stating that the workplace must be "permeated with discriminatory intimidation, ridicule and insult."[151] These results support Sperino, Thomas, and Beiner's respective arguments that the precedential force of extreme summary judgment decisions can cast a long shadow over subsequent cases.[152]

2. *Distinguishing Ordinary Work Conduct from Harassment*

Narrative responses also suggest that respondents were attempting to distinguish ordinary office conduct from a hostile work environment. Some respondents contrasted behavior they considered to be minor and behavior that crossed a line due to its seriousness and/or repeated character. One respondent noted, "The defendants *went too far with it*. The[n], one of them

---

[146] Yang v. Lakewood Mgmt., LLC, 918 F. Supp. 2d 1205, 1206 (D. Kan. 2013).

[147] *Id.* at 1208.

[148] Faragher v. City of Boca Raton, 524 U.S. 775, 788 (1998).

[149] *Yang*, 918 F. Supp. 2d at 1208 (discussing Morris v. City of Colorado Springs, 666 F.3d 654, 664 (10th Cir. 2012)).

[150] Taylor-Norman v. JoCo Assembly, No. 09-CV-410-TCK-FHM, 2010 WL 3521610 (N.D. Okla. Sept. 7, 2010).

[151] *Id.* at *7.

[152] SPERINO & THOMAS, *supra* note 50, at 37; Beiner, *supra* note 8, at 817.

ended up damaging her property. It was disgusting and unwarranted" (emphasis added).[153] Regarding the same case, another said, "[T]here are not simply one or two incidents. There are several. Even I would be uncomfortable with all the sex jokes; let alone giving customers my phone and home address."[154] Regarding a different case, a respondent explained: "the amount of sexual harassment and disturbing things that the defendant, Hanneman [sic] said about women and other comments are *way too far*" (emphasis added).[155] In another case, the respondent explicitly identified the behaviors he was willing to excuse versus those that crossed the line:

> I can understand a few jokes here and there, also the 'job is for a white girl' thing wasn't too bad. What did seem bad was: 1. The company made people sign a form saying the workplace was fine. 2. Standing over the man and badgering him to take a job was wrong. 3. The altered picture of a black man on his desk was wrong. 4. 'Get him out of the ghetto' statement seems wrong to me.[156]

Respondents seemed to be comparing the conduct in the fact pattern to their own assessment of the types of conduct that should or should not be permitted in the workplace. For example, one respondent characterized the facts in the case as: "inappropriate and crass sexual innuendo that should never be encouraged or tolerated in a professional environment."[157] Other respondents explained that "the use of racial slurs have no place in a workplace and no one should be subject to them,"[158] that "the utterance of her co-worker against her should never be tolerated by any employer,"[159] or that "the environment was too sexual . . . which made it pervasive."[160] Another observed that "[p]ornography and derogatory jokes do not belong in the workplace."[161] Similarly, another concluded, "[a]t a work place everyone should be professional and nothing physically of that nature should

---

[153] Respondent discussion of the facts in Strom v. Holiday Cos., 789 F. Supp. 2d 1060, 1071, 1086, 1090 (N.D. Iowa 2011) (denying summary judgment as to harassment claim).

[154] *Id.* at 1067–68, 1090.

[155] Respondent discussion of the facts in Henneman v. Airtran Airways, 705 F. Supp. 2d 1012, 1028–29, 1039 (E.D. Wis. 2010) (finding conduct sufficiently severe or pervasive to create an issue of fact, but granting summary judgment on other grounds).

[156] Respondent discussion of the facts in Green v. Harris Publications, Inc., 331 F. Supp. 2d 180, 183–84, 189, 191, 196 (S.D.N.Y. 2004) (granting motion to dismiss as to harassment claim).

[157] Respondent discussion of the facts in *Strom*, 789 F. Supp. 2d. at 1068–69 (discussing the facts therein).

[158] Respondent discussion of the facts in Austion v. City of Clarksville, 244 F. App'x 639, 646, 652, 654 (6th Cir. 2007) (affirming denial of motion of judgment as a matter of law as to harassment claim).

[159] Respondent discussion of the facts in Clay v. Credit Bureau Enters., Inc., 754 F.3d 535, 537–38, 541 (8th Cir. 2014) (affirming summary judgment as to harassment claim).

[160] Respondent discussion of the facts in Prince v. Madison Square Garden, 427 F. Supp. 2d 372, 375–76, 385 (S.D.N.Y. 2006) (denying motion to dismiss as to harassment claim).

[161] Respondent discussion of the facts in Rivera Abella v. Puerto Rico Tel. Co., 470 F. Supp. 2d 86, 106–07, 109 (D.P.R. 2007) (granting summary judgment as to harassment claim).

be done onto another person anywhere especially at work."[162] Several responses described the behavior in moralistic terms, characterizing it as "unacceptable behavior,"[163] "very much insulting,"[164] or "very toxic behaviors."[165] This language suggests that respondents' application of the severe or pervasive standard was informed by moral judgment—a form of assessment that judges are particularly ill-suited to perform on a jury's behalf when ruling on summary judgment.

The corresponding court rulings were generally devoid of moralistic language and instead attempted a more mechanical assessment of whether the conduct was frequent or serious enough to meet the threshold. In several of the cases referenced above, the court ultimately reached the same conclusion as respondents and denied summary judgment.[166] In others, however, courts appeared to lose sight of the forest for the trees, assigning little or no weight to certain types of evidence, and then declaring the remaining evidence insufficient. In two of the cases, the court applied the "stray remarks" doctrine—an evidentiary doctrine specific to causal inferences in disparate treatment discrimination claims that does not extend to harassment cases.[167] Courts also discounted or discarded evidence of pejorative comments that were made about other workers or about the plaintiff in their absence, ignoring the cumulative impact such comments might have on an employee's experience of the workplace. In some cases, the courts somehow managed to dismiss lengthy lists of harassing conduct—in one case declaring "thirty-plus" instances of racial harassment insufficient.[168]

For example, *Green v. Harris Publications* involved a Black plaintiff whose promotion was withdrawn after the hiring manager at a magazine was

---

[162] Respondent discussion of the facts in Jeffries v. Kansas, Dep't of Soc. & Rehab. Servs., 946 F. Supp. 1556, 1562, 1569 (D. Kan. 1996) (granting summary judgment as to harassment claim).

[163] Respondent discussion of the facts in Edwards v. Murphy-Brown, LLC, 760 F. Supp. 2d 607, 611–12, 633 (E.D. Va. 2011) (denying summary judgment).

[164] Respondent discussion of the facts in Lilly v. Roadway Express, Inc., 6 F. App'x 358, 358–59 (7th Cir. 2001) (affirming summary judgment).

[165] Respondent discussion of the facts in Clay v. Credit Bureau Enters., 754 F.3d 535, 537–38, 541 (8th Cir. 2014) (affirming summary judgment as to harassment claim).

[166] *See* Strom v. Holiday Companies, 789 F. Supp. 2d 1060, 1090 (N.D. Iowa 2011) (denying summary judgment); Henneman v. Airtran Airways, 705 F. Supp. 2d 1012, 1039 (E.D. Wis. 2010) (same); Austion v. City of Clarksville, 244 Fed. App'x 639, 654 (6th Cir. 2007) (same); *Edwards*, 760 F. Supp. 2d at 633 (same).

[167] *See Lilly*, 6 F. App'x at 359 (discussing the "stray remarks" doctrine); Green v. Harris Publications, Inc., 331 F. Supp. 2d 180, 192 (S.D.N.Y. 2004) (same); *see also* Price Waterhouse v. Hopkins, 490 U.S. 228, 251–252 (1989) (distinguishing "stray remarks" from evidence relevant to sex stereotyping); David M. Litman, *What is the Stray Remarks Doctrine? An Explanation and a Defense*, 65 CASE W. RSRV. L. REV. 823, 835 (2015) (discussing use of "stray remarks" doctrine in disparate treatment cases, defining "[s]tray remarks [as] those comments that, while perhaps discriminatory, do not truly show the discrimination was a motivating factor in the relevant employment decision").

[168] *Clay*, 754 F.3d at 540; *see also Green*, 331 F. Supp. 2d at 192–94 (granting the employer summary judgment despite multiple instances of racial discrimination); Taylor v. Chao, 516 F. Supp. 2d 128 (D.D.C. 2007) (finding five separate allegations were insufficiently severe and pervasive).

instructed by the publisher to "just hire a white girl."[169] Two thirds of respondents concluded the plaintiff met the severe or pervasive standard, including the respondent who was block-quoted in the excerpt above. However, the court granted summary judgment after assigning no weight to a variety of racist statements by the magazine's publisher, reasoning that the plaintiff had few individual interactions with the publisher himself.[170] The court also discounted other racist remarks or epithets not made to or about the plaintiff and ultimately based its conclusion on two lone pieces of evidence.[171] It was almost as if sifting through individual pieces of evidence distracted the court from the broader question of how a layperson might react to the evidence as a whole.

B. *The Number of Harassers and Victims*

Respondents also considered the number of harassers and the number of employees affected by the harassment in their overall assessment of whether the conduct was severe or pervasive.

Examples of such responses include the following (emphasis added):

- "She was clearly harassed with *multiple men* asking her very perverted questions and making sexual comments to her all the time."[172]
- "A pattern was established and an environment that normalized this behavior. *Multiple instances both from inmates and officers* were involved."[173]
- "It looks like *most of the white guys had ganged up* against him. Wearing clothes that had messages that we can say are meant to mock the black people and even suggesting that black people take long to study in school since they are dumber than any other race."[174]
- "One worker is being treated differently than the others, and *their coworkers* take their clothing, and dress up stuffed animals with their clothing, in a racially insulting manner. *The other workers* laughed and did an empathy display that suggests intention or agreement. The stuffed gorilla with Roger's shirt, is proof."[175]

---

[169] *Green*, 331 F. Supp. 2d at 183 (internal quotations omitted) (citation omitted).

[170] *Id.* at 193–94.

[171] *Id.* at 192–94.

[172] Respondent discussion of the facts in Morales v. GEO Grp., 824 F. Supp. 2d 836, 839–40, 852 (S.D. Ind. 2010) (denying summary judgment as to harassment claim).

[173] *Id.*

[174] Respondent discussion of the facts in Renfroe v. IAC Greencastle, LLC, 385 F. Supp. 3d 692, 698–99, 710 (S.D. Ind. 2019) (denying summary judgment as to harassment claim).

[175] Respondent discussion of the facts in Rogers v. City of New Britain, 189 F. Supp. 3d 345, 351, 360 (D. Conn. 2016) (denying summary judgment as to harassment claim).

Respondents also seemed to consider the harassment of multiple employees relevant to the severity or pervasiveness of the conduct, particularly when the harassment occurred in front of others. One respondent noted how a harasser "abuses every women in the restaurent [sic] with influence."[176] In another case, the respondent observed that the harasser made "*everyone feel uncomfortable* but *everyone noticed* it was way more toward the plaintiff" (emphasis added).[177] In a third case, the respondent noted "sixteen different examples of how the workplace environment treated *their employees* with abuse in a sexual manner" and that the conduct "involved *not only* [the plaintiff] *but others*" (emphasis added).[178]

It was somewhat unclear whether respondents believed that the harassment was more pervasive because it targeted multiple workers or whether the targeting of multiple workers further substantiated the plaintiff's claims. As Knapp and Heshizer previously found, cases involving witnesses tended to be more successful.[179] Indeed, some respondents made reference to the presence of witnesses as corroborating factors (e.g., "the largest piece of evidence is that there were witnesses to the harassing behavior"; "someone else testified stating that they saw it happen as well"; "Robinson got two other officers to report on what they had seen").

However, respondents' emphasis on the number of victims and perpetrators could also reflect their intuitions regarding how a reasonable person would experience harassment. As any victim of middle school bullying might attest, being targeted by multiple individuals would feel more oppressive than an environment with a single aggressor. Likewise, social contexts in which multiple people are being targeted suggests a more widespread problem than one involving a single individual.

The corresponding court rulings generally took the number of harassers into account and considered multiple harassers to be an aggravating factor in assessing the case.[180] However, the harassment of multiple employees did not consistently improve the plaintiff's case. In some instances, courts credited evidence that other employees had been harassed and denied summary judgment.[181]

---

[176] Respondent discussion of the facts in Benefield v. MStreet Ent., LLC, 197 F. Supp. 3d 990, 995–97, 1007 (M.D. Tenn. 2016) (granting summary judgment as to harassment claim).

[177] Respondent discussion of the facts in Jones v. Wichita State Univ., 528 F. Supp. 2d 1222, 1229, 1245 (D. Kan. 2007) (granting summary judgment as to harassment claim).

[178] Respondent discussion of the facts in Prince v. Madison Square Garden, 427 F. Supp. 2d 372, 375–76, 385 (S.D.N.Y. 2006) (denying motion to dismiss as to harassment claim).

[179] Knapp & Heshizer, *supra* note 53, at 113.

[180] *See* Morales v. GEO Grp., 824 F. Supp. 2d 836, 839–40, 852 (S.D. Ind. 2010) (denying employer's motion for summary judgment in case where there was multiple harassers); Renfroe v. IAC Greencastle, LLC, 385 F. Supp. 3d 692, 703, 710 (S.D. Ind. 2019) (same); *Rogers*, 189 F. Supp. 3d at 351, 360 (same); *Prince*, 427 F. Supp. 2d at 380, 385 (same).

[181] *See, e.g.*, *Prince*, 427 F. Supp. 2d at 380, 385 (denying the employer's motion to dismiss due to multiple harassers).

However, in other cases, courts discounted or excluded evidence that other employees had been harassed.[182] For example, in *Benefield v. MStreet Entertainment*, the plaintiff alleged that a male supervisor targeted female chefs with "bullying" and "aggressive" behavior, "regularly commandeer[ing] the kitchens" when female chefs were working.[183] When the plaintiff complained, the employer dismissed her concerns as "emotional."[184] The plaintiff had firsthand testimony of sexist comments by the manager and also proffered statements by other workers. All nine of the respondents who reviewed the facts in *Benefield* concluded that the severe or pervasive standard was met. However, the court noted the supervisor's sexist comments were not directed at her and dismissed statements from others as "secondhand."[185] The court then concluded the plaintiff's claim was based only on "four inappropriate comments" and it was less serious than a handful of egregious cases from the Sixth Circuit where summary judgment was granted for the employer.[186]

C. *Interference with Job Tasks and Adverse Employment Actions*

In *Harris v. Forklift*, the Supreme Court advised that assessments of severity or pervasiveness should consider whether the conduct "unreasonably interferes with an employee's work performance."[187] Courts tend to interpret this language to refer to the plaintiff's psychological distress regarding the work environment, which indirectly affected their performance.[188]

As previously noted, the respondents in our study were not provided with the language from *Harris*. However, they tended to emphasize interference with an employee's work performance in their assessment. Unlike the courts, however, they tended to focus heavily on harassing conduct that interfered with job tasks or access to work-related resources.

Examples of such responses include the following (emphasis added):

- "[Collins] was denied *office supplies* in comparison to what white teachers received."[189]

---

[182] Benefield v. MStreet Ent., LLC, 197 F. Supp. 3d 990, 1003, 1007 (M.D. Tenn. 2016); *Jones*, 528 F. Supp. 2d at 1234–35 (ignoring evidence that other female police officers complained about harassment by coworkers); *see also* Duffy v. Dep't of State, 598 F. Supp. 2d 621, 628 (D. Del. 2009) (discounting evidence of harassment of other employees).

[183] *Benefield*, 197 F. Supp. 3d at 997 (internal quotation marks omitted).

[184] *Id.* (internal quotation marks omitted).

[185] *Id.* at 1003.

[186] *Id.*

[187] Harris v. Forklift Sys., Inc., 510 U.S. 17, 23 (1993).

[188] *Harris*, 510 U.S. at 21–23 ("tangible psychological injury" not required to prove Title VII harassment claim, but noting that "[a] discriminatorily abusive work environment, even one that does not seriously affect employees' psychological well-being, can and often will detract from employees' job performance"); Edwards v. Murphy-Brown, LLC, 760 F. Supp. 2d 607, 629 (E.D. Va. 2011) (interpreting "unreasonably interferes" to mean psychological distress).

[189] Respondent discussion of the facts in Collins v. Meike, 52 F. App'x 835, 836, 838 (7th Cir. 2002) (summary judgment affirmed).

- "Collins alleged that she had to teach in second-floor classrooms, she had to use the back stairway of the school, she *received half the teacher's aid assistance* that the white teachers received."[190]
- "The black women *did not get any help from the white employees* and also she hard [sic] to get medicated for her injury."[191]
- "The male programmers also received *training that wasn't offered to the females*, and it was also mentioned that the supervisors treated the males differently than the females."[192]
- "For one, Ms. Johnson was pitted against a caucasian woman by various means, making Ms. Johnson look incompetent when she had *less staff on hand*, while the caucasian employee had more staff."[193]
- "She was held to higher standard than white employees, and reprimanded for simple things. Also, her *workplace requests were ignored*."[194]
- "*Tampering with her work tools* could also cause her harm or even to lose her job."[195]
- "[S]*tripping Nguyen of work* in different forms makes it clear Nguyen is not wanted, or welcomed."[196]

These descriptions reflect a common-sense view of hostile work environments. A plaintiff on the receiving end of various forms of harassing conduct would reasonably perceive the environment as particularly hostile if the plaintiff has also been denied access to the resources necessary to perform their job. Indeed, respondents sometimes cited these more tangible forms of work interference alongside other types of harassment conduct (such as derogatory comments or jokes) suggesting that they viewed the evidence in combination.[197]

---

[190] *Id.*

[191] Respondent discussion of the facts in Taylor-Norman v. JoCo Assembly, No. 09-CV-410-TCK-FHM, 2010 WL 3521610 at *7–*8 (N.D. Okla. Sept. 7, 2010) (summary judgment granted).

[192] Respondent discussion of the facts in Duffy v. Dep't of State, 598 F. Supp. 2d 621, 625, 632 (D. Del. 2009) (summary judgment granted as to harassment claim).

[193] Respondent discussion of the facts in Johnson v. Hosp. Corp. of Am., 767 F. Supp. 2d 678, 686, 724 (W.D. La. 2011) (summary judgment granted).

[194] *Id.*

[195] Respondent discussion of the facts in Ladd v. Grand Trunk W.R.R., 552 F.3d 495, 497, 503 (6th Cir. 2009) (summary judgment affirmed).

[196] Respondent discussion of the facts in Nguyen v. Mabus, 895 F. Supp. 2d 158, 166–67, 191 (D.D.C. 2012) (summary judgment granted).

[197] For example, the reference to interference with a worker's tools also discussed verbal harassment, noting: "The comments from coworkers are very hurtful, including those where they are talking about other women. That will make other women feel unsafe." In the case where the respondent

By contrast, courts deciding the same cases tended to overlook the harasser's attempt to sabotage or interfere with the plaintiff's work. Instead, court opinions were characterized by a kind of tunnel vision, where actions that interfered with the plaintiff's ability to perform their jobs were categorized as relevant only to a discrimination claim, or only to the question of whether the plaintiff suffered a tangible employment action for the purpose of vicarious liability.

For example, in *Collins v. Meike*, the plaintiff alleged that she "had to teach in second-floor classrooms, she had to use the back stairway of the school, she received half the teacher's aid assistance that the white teachers received, and she was allotted fewer school supplies than the white teachers."[198] These allegations implicate the terms and conditions of Collins's employment even if each did not individually rise to the level of an adverse employment action. They are also functionally similar to the earliest harassment case, *Rogers v. EEOC*, brought by a Latina optometry worker who alleged harassment based in part on her employer's practice of segregating patients based on national origin.[199] Two-thirds of the respondents who assessed Collins's case found these disparate conditions to be hostile. The narrative responses quoted above emphasized the effect on Collins's ability to do her job. Yet the court discounted the impact on job duties, assessing the severe or pervasive element of the claim only based on a derogatory remark by a supervisor and the lack of supplies, rather than the work environment as a whole.[200]

Likewise, in *Duffy v. Department of State*, the plaintiff alleged that she and other female computer programmers were mistreated by a co-worker and supervisor, did not have access to the same training, and that she was excluded from programmer meetings.[201] The plaintiff also alleged that she was "fearful" of her supervisor who would frequently yell at her and at one point threatened to slap her.[202] In assessing the severity or pervasiveness of the claim, the court ignored the various workplace obstacles plaintiff alleged, claiming the allegations were based on her "self-serving" interrogatory answers.[203] It then focused exclusively on the substance of the verbal abuse by the supervisor and co-worker, concluding that they did not use any derogatory references to the plaintiff's gender.[204] The court failed to appreciate the connection between the verbal abuse and the exclusionary treatment, and how both operated together to produce a hostile work environment.

---

noted that the plaintiff had been "stripped" of duties, the respondent also referenced comments by others and that they had mocked the plaintiff's accent.

[198] Collins v. Meike, 52 F. App'x 835, 836 (7th Cir. 2002).

[199] Rogers v. Equal Emp. Opportunity Comm'n, 454 F.2d 234, 236 (5th Cir. 1971).

[200] *Collins*, 52 F. App'x at 837.

[201] Duffy v. Dep't of State, 598 F. Supp. 2d 621, 624 (D. Del. 2009).

[202] *Id.* at 623, 625 (internal quotation marks omitted).

[203] *Id.* at 629.

[204] *Id.* at 628.

Ladd v. Grand Truck Western Railroad further illustrates the extreme disconnect between how our simulated jurors and courts respond to the same fact patterns. Nearly 75% of the 15 respondents who reviewed the *Ladd* fact pattern concluded that it met the legal standard, whereas the court granted summary judgment for the employer. The plaintiff was the only Black woman working as a welder in a department that almost exclusively consisted of white men. She brought harassment claims based on race and sex for constant "sexually degrading commentary."[205] The only other woman in the department confirmed that they were subject to "unwelcome sexual comments on a daily basis."[206] Another employee also referred to her as a "black bitch."[207] Ladd also described several separate instances of work tampering: "her shears were taken apart . . . her oxygen tanks here taken" and "bolts were removed from her grinder"—creating a "dangerous" situation.[208] As one respondent observed, "[t]ampering with her work tools could also cause her harm or even to lose her job."

On appeal, the Sixth Circuit declared that Ladd's case was insufficiently severe or pervasive.[209] It explicitly considered the evidence of work tampering, yet declared the actions "still do not rise to the level of an actionable hostile work environment" because there was no physical contact, and she was only subjected to one sex- or race-based epithet.[210] The court acknowledged that the "sex-based comments" occurred daily, but that the severity of the conduct was unknown—suggesting that it was applying a severe *and* pervasive standard, a misapplication of the legal standard.[211] The court then concluded that the work tampering was not dangerous *enough* to be considered severe, an odd conclusion that was not shared by most respondents that reviewed the same fact pattern.

D. *The Harasser and Organization's Response to Complaints*

Respondents also considered both the harasser's and the organization's response to any complaints of harassment to be centrally important in their assessment of the severity or pervasiveness of the conduct.

Several respondents highlighted the harasser's failure to "learn" from their mistakes or to improve following a warning from management. Examples of these responses included statements such as (emphasis added):

---

[205] Ladd v. Grand Trunk W.R.R., 552 F.3d 495, 497 (6th Cir. 2009).

[206] *Id.* at 498 (internal quotation marks omitted).

[207] *Id.* (internal quotation marks omitted).

[208] *Id.*

[209] *Id.* at 502.

[210] *Id.* at 501.

[211] *Id.* at 501; *see also* Kleinschmidt, *supra* note 80 at 1130 (arguing that courts are moving towards a "severe *and* pervasive standard").

- "The fact that he did not *learn from his punishment* and continued to harrass [sic] women around the office."[212]
- "There were multiple reports and behaviors. It was not a single instance or something that was taken out of context. It's clear that Herrera was a pervert that *did not learn from his mistakes*."[213]
- "He has touched the plaintiff on several occasions and *continued to do so* when asked to stop."[214]
- "On multiple instances, the plaintiff Duffy has proved that Fred was 'particularly nasty', and harassed multiple female coworkers. *He was instructed to stop and yet still continued*. He even threated coworkers with violence."[215]
- "Messer made it quite clear that *he was not going to stop his harassment*, and the management was not shown to help."[216]

Respondents also interpreted the harasser's refusal to stop as relevant to their broader culpability and motive for the conduct. The fact that the conduct continued even after a warning suggested that it was intentional rather than inadvertent. As one respondent put it, "[i]t was not a single instance or something that was taken out of context." In a way, respondents considered the conduct more severe because it was willful—even defying corporate instructions to stop—and more pervasive because it continued following a complaint.

Respondents likewise considered the organization's response to the harassment in assessing whether the work environment was hostile. One respondent itemized a litany of sexualized conduct at work (including "peep holes" in the bathroom) and noted "[w]hen these concerns were brought up to management nothing was done about it."[217] In another case, the respondent observed that "even after she complaint [sic] they did not take any response."[218] In another, the respondent observed that "I felt that the disregard of complaints about harassment and discrimination that were never

---

[212] Respondent discussion of the facts in Engel v. Rapid City Sch. Dist., 506 F.3d 1118, 1121–22, 1128 (8th Cir. 2007) (summary judgment affirmed).

[213] *Id.*

[214] Respondent discussion of the facts in Prindle v. TNT Logistics of N. Am., 331 F. Supp. 2d 739, 744, 752 (W.D. Wis. 2004) (summary judgment denied).

[215] Respondent discussion of the facts in Duffy v. Dep't of State, 598 F. Supp. 2d 621, 623, 625, 632 (D. Del. 2009) (summary judgment granted).

[216] Respondent discussion of the facts in McKeown v. Dartmouth Bookstore, Inc., 975 F. Supp. 403, 404, 407 (D.N.H. 1997) (summary judgment granted).

[217] Respondent discussion of the facts in Edwards v. Murphy-Brown, LLC, 760 F. Supp. 2d 607, 616, 633 (E.D. Va. 2011) (summary judgment denied).

[218] Respondent discussion of the facts in Morales v. GEO Grp., Inc., 824 F. Supp. 2d 836, 839 (S.D. Ind. 2010).

followed up led to a hostile work environment."[219] Another noted the company's failure to impose discipline on the harasser: "No mention of disciplinary action taken against the co-worker who planted the noose, which in any other situation would be considered a death threat."[220] A few respondents also treated multiple complaints as supporting evidence that the harassment occurred and that it was serious ("the way she was treated when she went to supervisors and HR seems to confirm this took place and she was either ignored or subjected to more harassment."[221]; "All the complaints she made reguarding [sic] the treatment she was receving [sic] shows that she was severly harrassed [sic]."[222]).

Courts generally did not consider the company's failure to respond—or the harasser's continued misconduct—as an aggravating factor in assessing the severe or pervasive nature of the conduct.[223] Instead, those facts tended to be confined to considerations of vicarious liability and the applicability of the *Faragher/Ellerth* defense. A rare exception to the trend was *Johnson v. Hospital Corp.*, where the court devoted serious attention to whether the employer's response contributed to the harassment, primarily to address the employer's argument that "all of plaintiff's claims of harassment . . . are actually claims of retaliation."[224] In that case, much of the harassment consisted of ignoring the plaintiff's requests and complaints, limiting her access to work resources, and reprimanding her for trivial matters. The court ultimately concluded that the employer's response could be considered in the severe or pervasive analysis because they related to a "term, condition, or privilege of employment."[225]

As the respondents intuitively grasp, a work environment is made substantially worse by continued harassment following a complaint—it suggests that there is little or nothing the plaintiff can do to further put a stop to the harassment, making them all the more powerless and victimized.[226]

---

[219] Respondent discussion of the facts in Johnson v. Hosp. Corp. of Am., 767 F. Supp. 2d 678, 711, 715, 718 (W.D. La. 2011).

[220] Respondent discussion of the facts in Taylor-Norman v. Joco Assembly, No. 09-CV-410-TCK-FHM, 2010 WL 3521610, at *1–2 (N.D. Okla. Sept. 7, 2010) .

[221] Respondent discussion of the facts in *Morales*, 824 F. Supp. 2d at 839.

[222] *Id.* at 839–42, 845, 850.

[223] Although some cases nevertheless survived summary judgment, others did not. *See, e.g.*, Engel v. Rapid City Sch. Dist., 506 F.3d 1118, 1123 (8th Cir. 2007) (summary judgment reversed on appeal); Prindle v. TNT Logistics of N. Am., 331 F. Supp. 2d 739, 756 (W.D. Wis. 2004) (denying summary judgment on issue of fact on severe or pervasive element); McKeown v. Dartmouth Bookstore, Inc., 975 F. Supp. 403, 406–07 (D.N.H. 1997) (summary judgment granted in part, denied in part); Duffy v. Dep't of State, 598 F. Supp. 2d 621, 629–630 (D. Del. 2009) (summary judgment granted); Edwards v. Murphy-Brown, LLC, 760 F. Supp. 2d 607, 631 (E.D. Va. 2011) (issue of fact on severe or pervasive element); Morales v. GEO Grp., Inc., 824 F. Supp. 2d 836, 852 (S.D. Ind. 2010) (denying summary judgment because of issue of fact on severe or pervasive element).

[224] Johnson v. Hospital Corp. of Am., 767 F. Supp. 2d 678, 711 (W.D. La. 2011).

[225] *Id.* at 716.

[226] Research in the psychology literature refers to a concept known as "institutional betrayal" where

E. *The "Reasonable" Woman Standard*

As Part I.D recounted, a variety of scholars have argued that the "reasonable person" standard for assessing the hostility of a work environment should be replaced with a "reasonable woman" standard. In this study, respondents were given standard jury instructions referencing a severe or pervasive standard based on how a "reasonable person in the plaintiff's position" would experience the work environment. Despite the absence of a "reasonable woman" instruction, or its equivalent for other protected categories, many respondents appear to have given serious thought to how a person in that protected category would feel when confronted with the behavior. In other words, they interpreted the jury instructions to consider "a reasonable person in plaintiff's position" to include both the situational and identity-based meanings of the word "position."[227]

In particular, participants highlighted conduct that would not necessarily be offensive if directed at someone outside that protected category but would be particularly offensive to someone of that race, sex, or religion. With respect to a religious harassment claim, one respondent explained, "religion is a sacred thing and to have offensive remarks made about a coworker's religious preference is wrong and to have it done on multiple occasions means it has [sic] harassment."[228]

Respondents demonstrated sensitivity towards race-based conduct that would be particularly threatening to Black workers. In the context of a race-based harassment claim, one respondent observed, "I also found that the noose was used to make any black employees feel uncomfortable and threatened."[229] In assessing a different case, another respondent explained "Calling a black person a monkey or the n-word are both examples of racism and pretty clear ones at that. Experiencing that would make the work environment hostile."[230] Another highlighted that "their coworkers take their

---

"[i]nstitutions have the potential to put those who depend on them in harm's way, through acts of omission or commission" for example, through a "lack of action to disrupt abusive patterns" and "allow[] continued abuse . . . ." Maria-Ernestina Christl, Kim-Chi Tran Pham, Adi Rosenthal & Anne P. DePrince, *When Institutions Harm Those Who Depend on Them: A Scoping Review of Institutional Betrayal*, 25 TRAUMA, VIOLENCE & ABUSE 2797, 2797 (2024). Institutional betrayal has been associated with psychological symptoms. In the workplace context, an employer can contribute to the abusive quality of a workplace environment by failing to take meaningful action following a complaint.

[227] *See* discussion *supra* note 84 ("Some commentators have concluded that the reference to 'plaintiff's position' in the *Oncale* standard implicitly includes the plaintiff's identity.").

[228] Respondent discussion of the facts in Wheeles v. Nelson's Elec. Motor Serv., 559 F. Supp. 2d 1260, 1264, 1270, 1271, 1274 (M.D. Ala. 2008) (granting summary judgment as to harassment claim).

[229] Respondent discussion of the facts in Taylor-Norman v. Joco Assembly, No. 09-CV-410-TCK-FHM, 2010 WL 3521610, at *7 (N.D. Okla. Sept. 7, 2010).

[230] Respondent discussion of the facts in Rogers v. City of New Britain, 189 F. Supp. 3d 345, 350–51 (D. Conn. 2016).

clothing, and dress up stuffed animals with their clothing, in a racially insulting manner . . . The stuffed gorilla with Roger's shirt, is proof."[231]

In another case, the respondent concluded that "continuing to find bananas on the truck, and coworkers who wore Confederate t-shirts confronting Jones, which in my opinion constitutes 'pervasive' harassment."[232] Another respondent assessing the same case engaged in similar reasoning, explaining "verbal racism was bad enough, but the bananas were abusive and extreme. The confederate iconography was also probably not the best idea."[233] Both of these respondents appear to have placed considerable weight on the gravity of the racist messaging associated with leaving a banana in a Black co-worker's truck, suggesting that they were applying the standard of a "reasonable Black person" in that situation. Had they instead applied a generic "reasonable person" standard, they would not have attached any particular significance to a banana left in a truck—outside the particular racial context, leaving a banana for a co-worker is not particularly offensive.

In the context of sex-based harassment cases, respondents also appear to have made subtle contextual inferences about the physical space in which the harassment took place, and its significance for how a woman would experience the conduct. Respondents evaluated claims regarding physical proximity in combination with other allegations regarding offensive verbal conduct by the harasser.[234] If the harasser had a track record of offensive conduct, respondents considered the ambiguous proximity to be further evidence of harassment. One respondent explained, "because of the fact she was [sitting] down and the limited space between her and the other employee, [the perpetrator's] buttocks or genitals would end up near her face or the face of a coworker."[235] Another respondent also highlighted the role of physical proximity in their assessment, noting: "Personal space is important. His standing inappropriately close is a power play."[236]

Describing another case, a respondent described how the cubicle layout and the supervisor's job duties provided a pretext for the perpetrator to harass the plaintiff undetected:

> He used his daily contact with her, such as asking her about her tasks, to interact with her and get close to her. Since it was

---

[231] *Id.* at 351.

[232] Respondent discussion of the facts in Jones v. UPS Ground Freight, 683 F.3d 1283, 1289–91 (11th Cir. 2012).

[233] *Id.*

[234] Respondent discussion of the facts in Rivera Abella v. P.R. Tel. Co., 470 F. Supp. 2d 86, 106–07 (D.P.R. 2007) (involving instances of "obscene language" and sexual remarks); Wyninger v. New Venture Gear, Inc., 361 F.3d 965, 973 (7th Cir. 2004) (involving sexual remarks).

[235] Respondent discussion of the facts in *Rivera Abella*, 470 F. Supp. 2d at 107.

[236] *Id.*

> in a cubicle and he was the supervisor, it could easily be hidden from others and could go on for many months.[237]

In another case, the respondent highlighted the threatening nature of male co-workers locking the door to an office: "once she already expressed discomfort, they continued to ask her inappropriate questions. Also, locking the door after she's inside a room is extremely creepy and terrifying."[238]

Several respondents also referenced how they would feel if they were in the plaintiff's shoes, while taking into account the plaintiff's particular protected characteristics. One wrote, "While I read through the accusations by Rivera-Rivera, I had vivid imaginations of how terrible it must have been for to go through such."[239] Another wrote, "I'd be very uncomfortable working in an environment in which I was treated differently from everyone of a different ethnicity and sex. Being subjected to racial slurs, fewer opportunities, and insubordination makes for a hostile work environment."[240]

Courts were inconsistent in their treatment of the same set of facts. In some cases, they displayed a degree of sensitivity to context, for example, noting the "particular racial history" of the use of the n-word or "the exceptionally ugly history of depicting African Americans as apes."[241] In other cases, courts ignored the context—for example, dismissing a race-based harassment claim that involved a noose left at a different employee's workstation, because she was not the "target" and because no "supervisor[] . . . participated in the noose incident."[242]

Courts were also inconsistent in their treatment of context surrounding physical proximity in gender-based harassment claims. In *Wyninger v. New Venture Gear,* the court noted the "physically intimidating situation" inherent in a "woman locked in a small room with three larger men, snickering at her refusal to discuss oral sex."[243] By contrast, in *Abella v. Puerto Rico Telephone*, the court concluded that the harasser's decision to place his buttocks near the plaintiff's face raised "no inference" that "they were based on plaintiff's sex"[244] despite evidence the harasser viewed

---

[237] Respondent discussion of the facts in Ferguson v. Chi. Hous. Auth., 155 F. Supp. 2d 913, 915, 918–19 (N.D. Ill. 2001) (denying motion for summary judgment).

[238] Respondent discussion of the facts in Wyninger v. New Venture Gear, Inc., 361 F.3d 965, 975–76, 981 (7th Cir. 2004) (discussing facts potentially presented an issue of fact as to whether the harassment was severe or pervasive, but affirming summary judgment on the basis the defendant successfully established the *Faragher/Ellerth* defense).

[239] Respondent discussion of the facts in *Rivera Abella*, 470 F. Supp. 2d. at 106–07.

[240] Respondent discussion of the facts in Dragon v. Conn., 211 F. Supp. 3d 441, 444–46 (D. Conn. 2016).

[241] Rogers v. City of New Britain, 189 F. Supp. 3d 345, 355–56 (D. Conn. 2016); *see also* Jones v. UPS Ground Freight, 683 F.3d 1283, 1297–98 (11th Cir. 2012) (considering the history of racial stereotypes against African Americans).

[242] Taylor-Norman v. Joco Assembly, No. 09-CV-410-TCK-FHM, 2010 WL 3521610, at *7 (N.D. Okla. Sept. 7, 2010).

[243] *Wyninger*, 361 F.3d at 977, 981 (affirming summary judgment as to harassment claim based on *Faragher* defense).

[244] *Rivera Abella*, 470 F. Supp. 2d at 107.

pornography at work, "use[d] . . . obscene language and jokes of a sexual nature," and stared at the plaintiff "through a hole in the distribution frame while she worked at her desk."[245]

These results suggest that respondents applied a standard approximating a "reasonable woman" standard, and in some cases a "reasonable Black person," based on the standard jury instruction to consider how a "reasonable person in plaintiff's position" would experience the "work environment." Courts, by contrast, were quite inconsistent with respect to whether they considered the social and historical context in evaluating whether the severe or pervasive standard was met.

F. *Conduct that Respondents Concluded Did Not Meet the Severe or Pervasive Standard*

As illustrated in Table 3, the survey results suggest that 67 of the cases in the sample should have survived summary judgment, while 8 cases should have been dismissed. As previously discussed, these unsuccessful cases included a disproportionate number of race-based claims, and intersectional claims. Recognizing the possibility that these outcomes were driven at least in part by respondent bias, we nevertheless attempt to distill the narrative responses provided for these cases to identify the types of facts that led respondents to conclude that the severe or pervasive standard was not met. Below, we separate the 8 cases into 3 groups, summarize the facts of the case, the narrative responses associated with each group, and offer our recommendation regarding whether cases falling within that category should properly be dismissed on summary judgment.

The 8 cases that did not meet the our measurement threshold for what a "reasonable jury" could find severe or pervasive standard fell into 3 categories: (1) hostile or rude behavior in connection with personnel actions; (2) claims based on a single incident that was innocuous, minor, or not at all related to the plaintiff's protected category; and (3) a claim alluding to racial epithets outside of the plaintiff's presence. Each category is discussed below.

1. *Hostile or Rude Behavior in Connection with Personnel Actions.*

The cases falling into the first category were *Atanus v. Perry*,[246] *Brooks v. Grundmann*,[247] and *Baird v. Snowbarger*.[248] All three plaintiffs were

---

[245] *Id*. at 106–07.

[246] Atanus v. Perry, 520 F.3d 662, 667–69 (7th Cir. 2008).

[247] Brooks v. Grundmann, 748 F.3d 1273, 1275–76 (D.C. Cir. 2014).

[248] Baird v. Snowbarger, 744 F. Supp. 2d 279, 283–85, 296 (D.D.C. 2010). The case was reversed in part on appeal (following the collection of the sample), on the basis that the district court improperly excluded evidence. Baird v. Gotbaum, 662 F.3d 1246, 1253 (D.C. Cir. 2011). The study only included the facts set forth at the district court level, and the outcome was coded based on the district court dismissal.

female government employees, Atanus was white, and both Brooks and Baird were Black.

Brooks alleged that her supervisor "yell[ed] at her in front of co-workers" and "fl[ung] a heavy notebook which Brooks thought was aimed in her direction."[249] Atanus alleged that her supervisor "badgered her in a 'loud, unprofessional tone.'"[250] Both Brooks and Atanus received performance counseling—Atanus received multiple "letter[s] of instruction" documenting her performance deficits.[251]

Baird v. Snowbarger is a somewhat idiosyncratic case, where the plaintiff alleged harassment on the basis race and sex, as well as retaliatory harassment for her protected activity.[252] The facts in *Baird* were also somewhat complex because they occurred in connection with union activity, an internal investigation into Baird's union activity, and an arbitration. Baird—herself an attorney—alleged that a coworker referred to her as "psychotic" in an email; that an adverse attorney in a union arbitration warned others to watch what they say because certain staff (allegedly referring to plaintiff) would "twist[] and publiciz[e]" their words; and that an adverse witness in a deposition "began angrily yelling at plaintiff and pounding the table."[253]

Respondents evaluating these cases observed both the lack of a connection between the alleged behavior and the plaintiff's protected category, and the more plausible connection between the behavior and the fraught interpersonal context of performance counseling. In assessing *Brooks*, one respondent wrote, "While I wouldn't condone the way her boss acted I don't believe it would qualify as a violation. I also don't see any evidence it was race based." Other respondents made similar comments, noting "it just sounded as if her supervisor was unhappy with the work that Mrs. Brooks had done" or that "it could have been totally based on performance and errors." In assessing *Atanus*, respondents emphasized the lack of "detail" and "information" about her allegations, her poor performance, and the ordinary nature of the alleged conduct: "All the plaintiff claimed was that she was yelled at. It could have happened to anyone." Several noted there was "nothing," "next to nothing," or "no mention" of Atanus's protected category (color, race, religion—depending on the fact pattern the respondent received).

Likewise, the hostile behavior in *Baird* occurred in the context of workplace conflicts over plaintiff's alleged dissemination of union-related flyers, investigations regarding the flyers, related HR admonitions, and further investigations. Respondents noted the likely connection between

---

[249] *Brooks*, 748 F.3d at 1275.
[250] *Atanus*, 520 F.3d at 669 (citation omitted).
[251] *Brooks*, 748 F.3d at 1275; *Atanus*, 520 F.3d at 669–70.
[252] *Baird*, 744 F. Supp. 2d at 294.
[253] *Id.* at 284–85, 294.

these conflicts and adversarial behavior towards Baird, and the absence of evidence that they were attributable to the plaintiff's sex or race: "There wasn't any information presented that indicated any harassment at all that could be based on sex. There was animosity in the office towards Baird but that was based on the perception that Baird was spreading false accusations through the workplace."[254]

The narrative responses suggest that respondents were making inferences and drawing conclusions based on the context, where supervisors were frustrated with poor performance or coworkers were aggravated by multiple investigations. While the latter might qualify as retaliation, the presence of those alternate explanations for the hostile treatment appears to have undermined the potential inference that the conduct related to their protected category. In other words, the same kind of contextual reasoning that led respondents to recognize potential harassment in other circumstances led them to rule it out in *Atanus*, *Brooks*, and *Baird*.

Thus, we conclude these types of Title VII harassment claims may be properly dismissed on summary judgment, if there is no evidence that the conduct was based on a protected category. We also caution that this category should be limited to circumstances where there is a clear alternative explanation for the hostile behavior. Note that the alternative explanation may give rise to a different legal claim, such as Title VII retaliation or a Section 8(a)(1) violation of the NLRA.

2. *A Single Incident That was Relatively Innocuous.*

The cases falling into the second category included *Saunders v. Emory Healthcare*,[255] *Nurriddin v. Goldin*,[256] *Torres v. City of Philadelphia*,[257] and *McCray v. Wal-Mart*.[258] Those cases involved a single instance of conduct that was innocuous or minor, including being asked not to wear a beret to work per company policy (*McCray*); being falsely accused of egging a co-worker's car (*Torres*); noticing empty digital folders on a co-worker's computer labeled "'racist jokes and stories'" and "'sex bulletin board'" (*Nurridin*); posting "motivational signs" in "employee areas" that may or may not have intended to "depict [B]lack nurses" (*Saunders*).[259]

Respondents noted the lack of detail, the isolated nature of the event, and the innocuous or minor nature of the conduct. For example, respondents wrote: "there was not much detail at all in what I read to be honest," "who

[254] Some respondents were directed to assess Baird's race-based case, while others were instructed to assess Baird's sex-based case.

[255] Saunders v. Emory Healthcare, Inc., 360 F. App'x 110, 117 (11th Cir. 2010).

[256] Nurriddin v. Goldin, 382 F. Supp. 2d 79, 109 (D.D.C. 2005).

[257] Torres v. City of Philadelphia, 907 F. Supp. 2d 681, 689 (E.D. Pa. 2012).

[258] McCray v. Wal-Mart Stores, Inc., 377 F. App'x 921, 925 (11th Cir. 2010).

[259] McCray v. Wal-Mart Stores, Inc., No. 1:06-CV-1123-MEF, 2009 WL 734138 at *6 (M.D. Ala., Mar. 17, 2009); *Torres*, 907 F. Supp. 2d at 685; *Nurridin*, 382 F. Supp. 2d at 87; *Saunders*, 360 F. App'x 110 at 117.

gives a shit?" "it was a one time incident." Respondents also commonly noted the absence of a connection between the conduct and the plaintiff's protected category. In the *Torres* case based on a false accusation of egging a car, respondents noted the absence of evidence that the accusation was "race-motivated": for example, saying "I don't know what [accuser] chose Torrez [sic] and [another officer] to blame. There could be any number of circumstances, histories, and other factors."[260] With respect to the *Nurriddin* case, many respondents emphasized that the folders on the computer were empty, suggesting perhaps that they might have reached a different conclusion had the folders contained offensive material.[261]

The cases falling into this category can properly be dismissed on summary judgment and can be readily identified by courts—because they notably (1) involve a single incident, *AND* (2) the incident itself is objectively minor or harmless.

3. *Racial Epithets Uttered in the Plaintiff's Presence.*

The final case that did not meet the 40% threshold was *Burkett v. Glickman*. In *Burkett v. Glickman*, the plaintiff alleged that she was "present during the utterance of racially prejudicial remarks by supervisors."[262] Burkett apparently provided no detail regarding these allegations other than deposition testimony from a co-worker who said he heard their supervisor use the n-word "in front of certain . . . employees."[263] Few respondents provided detailed commentary regarding their conclusion that the standard was not met. One explained that the *Burkett* case involved "very little evidence" because "only one employee testified that he had heard this."

We believe that cases involving racial epithets are unsuitable for summary judgment, as they can be difficult to distinguish from other cases that jurors believed were severe based on a single incident. Although such cases may ultimately fail before a jury, it is properly the jury's role to assess the severity of such incidents, rather than leaving courts to anticipate which types of severe but infrequent conduct pass the threshold, and which do not.

G. *Limitations*

Our methodology is subject to a few potential limitations and caveats. First, the summary judgment rate we observed for harassment cases is specific to a subset of harassment cases, not harassment cases generally. Our Westlaw-based sampling technique does not include, for example, cases settled prior to a motion for summary judgment. This could inflate the rates of summary judgment observed in this study.

---

[260] Respondent discussion of the facts in *Torres*, 907 F. Supp. 2d at 685.
[261] Respondent discussion of the facts in *Nurriddin*, 382 F. Supp. 2d at 86–88.
[262] Burkett v. Glickman, 327 F.3d 658, 661 (8th Cir. 2003).
[263] *Id.*

Second, our respondent sample may not reflect how juries would actually behave if these cases went to trial. In addition to the demographic differences between our sample and national demographics, jury selection is local in nature, and juries selected in any given location may differ significantly from our respondents in their assessments of the facts.[264] Our respondents may also assess facts differently from how jurors may have responded to facts decades ago, which could overestimate the differences between how courts and prospective jurors respond to facts. Our respondents also did not have the opportunity to deliberate, and it is possible the group dynamics of deliberation could make juries more or less sympathetic to plaintiffs.

V. IMPLICATIONS

Broadly, the results of this study suggest that courts should be far more restrained in dismissing harassment cases on summary judgment when the defendant challenges the conduct as insufficiently "severe or pervasive" to qualify as a hostile work environment. Courts dismissed about two-thirds of the cases in our sample on summary judgment, while our study suggests that number should be closer to 10%−17% (see Table 3).

Our results offer insights that can guide how judges should assess motions for summary judgment in harassment cases. Given the substantial disparity between how courts and our simulated jurors evaluated the facts, courts should recalibrate their stance on summary judgment. Courts should refrain from granting summary judgment unless the plaintiff presents next to no evidence supporting their case or where a plaintiff's meagre case would not withstand a mountain of evidence from the defendant. As our results show, that standard is rarely met in Title VII harassment cases where the severe or pervasive standard is the basis for a summary judgment motion.

Our quantitative results also suggest some possible explanations for why courts and respondents differ so starkly in their assessments. They also allow us to exclude certain alternate hypotheses that might appear to be driving the difference. First, the difference does not appear to be that judges are shifting their assessment of harassment cases over time, nor that the alleged harassment has become more or less severe or pervasive over time. The non-significant results reported in Tables 4 and 5 above rule out a time-based explanation for the divergence.

---

[264] National census-level statistics on juror demographics are not available, and most state-level statistics, where available, only report on racial representation. A comprehensive recent study of juror demographics in Washington State found that Black and Native Americans are underrepresented among those reporting for jury service, while those with higher levels of education and older Americans were over-represented. Peter Collins, Brooke Gialopsos and Bailey Tanaka, *Statewide Juror Summons Demographic Survey Project: An Analysis of Selected County Data* – 2023 Final Report, WASHINGTON STATE COURT vii, 28, 25 (June 30, 2023) https://www.courts.wa.gov/subsite/mjc/docs/Statewide%20Juror%20Summons%20Demographic%20Survey%20Project%202023.pdf.

Second, the results largely do not support the hypothesis that demographic differences between judges and the general population lead them to reach different conclusions on harassment claims of this sort. Gender did not produce significantly different ratings on our survey (see Table 7), and neither did political views. Although higher levels of respondent education and older age was statistically associated with higher ratings, the results were not in a direction that would account for the differences in case outcomes. Because all judges have graduate degrees, they have a much higher average level of education than the respondents in our study and the U.S. population generally. However, more education was associated with higher severe or pervasive ratings and therefore would not explain why judges are more likely to dismiss harassment claims than our simulated jurors. The same was true of age—judges are much older than the general population.[265] Yet results suggest that respondents over age 65 were more likely to give higher severe or pervasive ratings than younger respondents. This too fails to account for the divergence in outcomes between judges and our simulated jurors.

Black respondents were more likely to give higher severity ratings to fact patterns than respondents of other races. However, this difference did not drive the results overall—Black respondents represented only a small part of our sample (16%), and the respondents in our survey quite consistently reached different conclusions than judges. In addition, Black judges are currently proportionally represented within the federal judiciary—constituting 12% of federal judges, compared to 12% of the population. [266] The problem then cannot be attributed to a mismatch between the racial makeup of the judiciary and the population as a whole.

The results support the hypothesis that certain types of cases fare more poorly on summary judgment than others. As reported in Table 6, intersectional cases based on multiple protected categories were much less likely to succeed on summary judgment, even when controlling for severe or pervasive ratings. The disadvantage for intersectional cases may flow from the structure of harassment claims, where each protected category serves as a separate cause of action. This artificial separation means that each instance of harassment would typically be attributed to one protected category or another—effectively whittling down the available evidence for each claim. Consequently, evidence that would be severe or pervasive in the aggregate might be insufficient once parceled out to the various protected categories that formed the basis for the claim. This outcome places intersectional plaintiffs at a substantial disadvantage. And an employee that

---

[265] *Demography of Article III Judges, 1789-2020*, FED. JUD. CTR., https://web.archive.org/web/20210214143325/https://www.fjc.gov/history/exhibits/graphs-and-maps/age-and-experience-judges#expand (last visited Oct. 25, 2025).

[266] AM. BAR ASS'N, PROFILE OF THE LEGAL PROFESSION 57 (2023), https://www.americanbar.org/ content/dam/aba/administrative/news/2023/potlp-2023.pdf.

experiences harm based on multiple protected categories suffers the same harm from a hostile work environment as one that has been targeted for a single protected category.

The qualitative results—particularly when combined with the court rulings themselves—offer a detailed basis for understanding why our simulated jurors reached different conclusions than the courts. Put simply, the jurors tended to assess the evidence as a whole, while judges tended to "slice and dice" the evidence, as so many scholars observed. Jurors found that the presence of multiple harassers or multiple victims was an aggravating factor in whether the workplace was hostile. While courts found the presence of multiple harassers significant, they tended to discount evidence that other employees were subjected to harassment. Jurors also focused heavily on the harasser's culpability for continued conduct following a warning from the company, as well as the company's complicity for failing to stop the harassment. By contrast, courts tended to ignore such facts in assessing whether the conduct was severe or pervasive, treating them as relevant only to matters of vicarious liability.

In many ways, the simulated jurors do a better job than courts in applying the Supreme Court's guidance in *Harris v. Forklift* to consider "all the circumstances."[267] Respondents were also far more effective than courts in interpreting the facts in a manner consistent with the text of Title VII, which requires that the harassment alter the "terms [and] conditions . . . of employment."[268] Their focus, for example, on facts suggesting that the harasser was interfering with the employee's work tasks, resources, or equipment, is much closer to a textualist interpretation of the statute.

Furthermore, as Theresa Beiner argued, jurors are better suited to apply "community standards of appropriate behavior in the workplace" that distinguishes ordinary workplace conduct from harassment.[269] Respondents' conclusions were supported by reasoning regarding whether the conduct "went too far," whether the harasser "learned their lesson," or whether the conduct was "inappropriate," or "toxic." They also distinguished more serious conduct from conduct they "can understand," which could be explained by "other factors," lacked "detail" or was "a one time incident."[270] And in combining their assessment of the facts with their moral judgments, they reached starkly different conclusions than courts weighing the same evidence.

The results of this study suggest that the problem lies not with the severe or pervasive standard itself, which jurors appear to understand quite well based on their narrative responses. Rather, judges appear to vastly overestimate the amount of evidence a jury would consider sufficient to

---

[267] Harris v. Forklift Sys., Inc., 510 U.S. 17, 23 (1993).
[268] 42 U.S.C. § 2000e-2(a)(1).
[269] Beiner, *supra* note 8, at 820.
[270] *See supra* Part IV.

satisfy the standard. Similarly, the results of the study do not suggest that the "reasonable person" standard is problematic or disadvantageous to plaintiffs when presented to simulated jurors. Based on their narrative responses, jurors already appear to be applying a version of the "reasonable woman" standard, which takes into account how a person's perception of conduct would be shaped by their membership in a protected category.

These results also raise the counterfactual question of whether these cases would have ultimately prevailed had they proceeded to a jury. As previously discussed, David Oppenheimer's 2003 study of jury verdicts suggested that surviving summary judgment is no guarantee of success. Sex-based harassment cases fared the best—at a 68% win rate, while race-based harassment claims won in half of cases. [271] If more cases survive summary judgment, many of them may not ultimately succeed at trial. Even so, permitting plaintiffs to proceed to trial serves the plaintiff's interest in procedural justice: having their day in court. Furthermore, the results of our study suggest that juries are well-suited to weigh the evidence. As Beiner argued more than 20 years ago, courts should "let the jury decide" harassment cases.[272]

Courts should exercise far greater caution in granting summary judgment in harassment cases. They should begin from the presumption that harassment cases should almost never be dismissed because the court considers the facts insufficiently severe or pervasive. Our results suggest the only fact patterns juries would consistently reject as insufficiently severe or pervasive are those involving (1) a single minor or innocuous incident; or (2) where there was a clear alternate explanation for the situation, such as conflict in connection with performance counseling.

In making this assessment, courts should aggregate all of the evidence together, including evidence regarding the company's and the harasser's response to a complaint, evidence regarding work-related obstacles, and conduct involving harassers or victims. Avoiding the judicial tendency to "slice and dice" will protect against mistaking a viable harassment claim with a thin claim with virtually no supporting evidence. A fulsome consideration of all the evidence as to the severe or pervasive nature of the work environment does not preclude the court from later focusing on certain elements particularly relevant to the *Faragher/Ellerth* defense.

Courts should guard against relying on precedent involving extreme facts where a district or appellate court granted summary judgment.[273]

---

[271] Oppenheimer, *supra* note 66, at 535, 540, 543.

[272] *See generally* Beiner, *supra* note 8.

[273] The Sixth Circuit seems particularly inclined to produce bad precedent and follow bad precedent. *See, e.g.*, Ladd v. Grand Trunk W.R.R., Inc., 552 F.3d 495, 498, 501–03 (6th Cir. 2009) (stating that many instances of sex- or race-based comments and equipment tampering was insufficient and subsequently affirmed summary judgment); Benefield v. MStreet Ent. LLC., 197 F. Supp. 3d 990, 1003 (M.D. Tenn. 2016) (granting summary judgment); Goller v. Ohio Dep't of Rehab. & Corr., 285 F. App'x.

Extreme cases seem to have had an outsize influence on later summary judgment rulings, yet our study suggests that juries would be highly unlikely to reach the same conclusion. The true metric for assessing harassment claims should be the Supreme Court's guidance in *Harris v. Forklift* to consider "all the circumstances."[274]

Second, courts should consider intersectional harassment as a "reasonably comparable evil" to single-basis harassment and as such, within the "broad and flexible" interpretation the Supreme Court applied to Title VII's "because of" language.[275] Through this method, plaintiffs would need not prove that one or the other protected category motivated the harassment, but rather that their membership in the protected categories motivated the harassment. The plaintiff, after all, is a single individual. As long as the harassment was experienced due to membership in one or both protected groups, it should not matter which of those categories motivated each act of harassing conduct. Indeed, even the harasser themselves may not be aware of or fully able to articulate whether the harassment was motivated by one or both protected categories. Using the existing Supreme Court precedent of *Oncale v. Sundowner Offshore Services, Inc.* to present intersectional harassment as a reasonably comparable evil to single-basis harassment may be particularly effective at overcoming judicial skepticism.[276]

CONCLUSION

This study sought to empirically assess whether courts are correctly identifying cases with no genuine issue of material fact on summary judgment. Our study focuses on summary judgment cases in the Title VII harassment context, where defendants frequently challenge harassment as insufficiently "severe or pervasive" to qualify as a hostile work environment. We present the facts of actual cases to simulated jurors through a Qualtrics survey, along with jury instructions, and ask them to evaluate whether the plaintiff met the severe or pervasive standard.

To compare court outcomes with our survey responses, we categorize a case as not suitable for summary judgment if at least 40% of respondents thought the severe or pervasive standard was met. Based on this threshold, 89% of the cases in the sample should survive summary judgment, and only 11% should be dismissed. This result was in stark contrast to the actual results of the cases, where 65% of the cases in the sample were actually dismissed and only 35% survived.

Our survey also included qualitative data—specifically, respondents' unstructured response to a question asking them to explain the basis for their

250, 259 (6th Cir. 2008) (stating the fact that plaintiff was called "Barbie" over thirty times was infrequent and affirming summary judgment).

[274] Harris v. Forklift Sys., Inc., 510 U.S. 17, 23 (1993).

[275] Berning-O'Neill, *supra* note 9, at 908 (internal quotation marks omitted).

[276] *Id.* (internal quotations omitted).

decision. Their qualitative responses suggest respondents assess the evidence in a far more holistic manner than courts, which tend to "slice and dice" evidence based on their relevance to specific elements or defenses, rather than considering how an employee belonging to that protected category would experience the workplace.

Surprisingly, our simulated jurors analyzed facts in a manner far more consistent with the Supreme Court's guidance in *Harris v. Forklift* than the courts did. Unlike many of the court rulings included in the sample, jurors considered "all the circumstances" and placed particular weight on conduct affecting the terms and conditions of employment. This result further supports the conclusion that the jurors were faithfully applying the jury instructions and that courts have been far too aggressive in granting summary judgment in Title VII harassment claims.

Moving forward, judges should exercise extreme caution when defendants move for summary judgment on the basis that the conduct was insufficiently severe or pervasive to qualify as harassment. Courts should also allow intersectional harassment claims to be pled as a single cause of action to avoid prejudicing these claimants.