DOES THE PRESUMPTION OF VALIDITY MATTER? AN EXPERIMENTAL ASSESSMENT

Jeremy W. Bock *

INTRODUCTION

"*We don't know exactly how often the presumption makes a difference to a case outcome.*"[1]

In patent law, the presumption of validity[2] exerts a profound influence on litigation strategy.[3] It has attracted criticism—not only from academics[4] but also from at least one federal judge[5]—

 Assistant Professor of Law, University of Memphis Cecil C. Humphreys School of Law. The author thanks Bernard Chao, John Golden, Richard Gruner, Parisa Jorjani, Mark Lemley, Su Li, Evelyn Mak, Shawn Miller, Lisa Larrimore Ouellette, Efthimios Parasidis, Greg Reilly, David Schwartz, Ryan Vacca, Michelle Yang, his colleagues at Memphis Law, and the participants at the June 2013 Workshop on Research Design for Causal Inference at Northwestern University School of Law, the Fifth Annual Legal Scholars Conference at Arizona State University College of Law, PatCon4 at University of San Diego School of Law, the 2014 SEALS Conference, and the 14th Annual Intellectual Property Scholars Conference at Berkeley Law for helpful discussions and feedback. The author thanks the University of Memphis Cecil C. Humphreys School of Law for supporting this research. The survey experiment reported in this article was conducted pursuant to a protocol approved by the Institutional Review Board at the University of Memphis.

1. Doug Lichtman & Mark A. Lemley, *Rethinking Patent Law's Presumption of Validity*, 60 STAN. L. REV. 45, 70 (2007).

2. 35 U.S.C. § 282(a) (2012) ("A patent shall be presumed valid.").

3. For example, accused infringers may prioritize noninfringement defenses over invalidity defenses because of the heightened burden associated with proving invalidity. *See, e.g.*, Roger Allan Ford, *Patent Invalidity Versus Noninfringement*, 99 CORNELL L. REV. 71, 118 (2013) (observing that "the elevated burden of proof that applies to invalidity, . . . which stems from the statutory presumption that a patent is valid unless proved otherwise, makes it relatively more difficult to win an invalidity defense than a noninfringement defense even if the two defenses would otherwise have similar merits").

4. *See, e.g.*, F. Scott Kieff, *The Case for Preferring Patent-Validity Litigation Over Second-Window Review and Gold-Plated Patents: When One Size Doesn't Fit All, How Could Two Do the Trick?*, 157 U. PA. L. REV. 1937, 1955 (2009) ("[A] weakening of the presumption of validity would be particularly good for the 'Davids' of the system who face off against the 'Goliaths.' It directly protects them from the in terrorem effect of junk patents . . . ."); Lichtman & Lemley, *supra* note 1, at 47 ("[T]he law makes [patent]

for making weak patents difficult to invalidate. When mentioned to the jury, the presumption is perceived by litigants as exerting a powerful pro-patentee influence that overshadows its nominal procedural function of assigning the burden of proving invalidity.[6]

Despite its apparent strategic importance in patent litigation, hardly any empirical studies exist on whether and to what extent the presumption may affect how jurors decide invalidity issues, thereby leaving many basic questions unanswered and unexamined in the academic literature. For example, would mentioning the presumption to the jury actually affect the likelihood that it will find a patent invalid? If so, to what degree? Because the presumption assumes a level of administrative correctness,[7] should the jury ever be informed of the operational deficiencies (e.g., application backlog, quality of examiner review) of the U.S. Patent & Trademark Office ("PTO")?[8] If so, would such information undermine the presumption of validity?

To help answer these questions, this article reports the results of the first experimental study on the effect of instructing the jury on the presumption of validity. The impact on case outcomes when the presumption is mentioned, and whether criticisms about the PTO might counteract its influence (or vice versa), have long been the province of speculation. Experimental analysis may provide additional insights that could help refine intuitions about

issuance mistakes hard to reverse. The culprit is a legal doctrine known as the *presumption of validity*.").

5. *See* William Alsup, Memo to Congress, *A District Judge's Proposal for Patent Reform: Revisiting the Clear and Convincing Standard and Calibrating Deference to the Strength of the Examination*, 24 BERKELEY TECH. L.J. 1647, 1648 (2009) ("A central reason for the litigation boom is the presumption of validity and the 'clear and convincing' standard. . . . This presumption of validity applies equally to all patents—even those that are almost certainly invalid. This is a huge advantage for the patent holder—and it is often an unfair advantage . . . .").

6. *See, e.g.*, William G. Childs, *The Implementation of FDA Determinations in Litigation: Why Do We Defer to the PTO but Not to the FDA?*, 5 MINN. INTELL. PROP. REV. 155, 172 (2004) ("The psychological impact of this presumption of validity is difficult to measure. However, it is not insignificant that a jury is instructed by the one nominally neutral person in the courtroom that it must begin deliberations with the belief that the patent is valid." (emphasis omitted)).

7. Applied Materials, Inc. v. Advanced Semiconductor Materials Am., Inc., 98 F.3d 1563, 1569 (Fed. Cir. 1996) ("The presumption of validity is based on the presumption of administrative correctness of actions of the agency charged with examination of patentability." (citing Interconnect Planning Corp. v. Feil, 774 F.2d 1132, 1139 (Fed. Cir. 1985))).

8. *See infra* notes 65–67.

whether and when a presumption instruction may have effects on the case that go beyond the nominal procedural role contemplated by the Federal Circuit, which views the presumption instruction as an optional feature so long as the jury is informed of the clear and convincing standard of proof for invalidity.[9] In the Federal Circuit's view, any significance carried by the presumption instruction is subsumed in the clear and convincing standard.[10] This is contrary to the view commonly held among litigants that jurors are profoundly influenced by the presumption instruction, separate and apart from the standard of proof.[11] The potential consequences of this mismatch in perceptions have largely escaped scholarly attention.

The results of the experiment reported in this article largely confirm the conventional view of litigants that the presumption instruction may have a substantial impact on the jury's decision on invalidity issues. The data reveal statistically significant[12] differences in the rate at which mock jurors found invalidity based on whether they were informed of the presumption. Based on this finding, the potential exists for forum shopping arising from the absence of a consensus among trial judges on whether the presumption instruction should be included.

Part I of this article provides background information on the presumption of validity that is relevant to the experimental study. Part II describes the methodological design of a survey experiment used to test the conventional assumptions regarding the effect of mentioning the presumption and criticisms about the PTO during trial. Part III reports the results of the survey experiment. Based on the data, Part IV analyzes the procedural considerations and the error costs associated with the presence or absence of the presumption instruction. Part V discusses the limitations of this study, and is followed by a brief conclusion.

9. Chiron Corp. v. Genentech, Inc., 363 F.3d 1247, 1258–59 (Fed. Cir. 2004).

10. *Id*.

11. *See infra* note 17 and accompanying text.

12. In this article, a *p*-value less than 0.05, which corresponds to significance at the 5% level, will be treated as the threshold for statistical significance. Where appropriate, *p*-values less than 0.10 but greater than or equal to 0.05, which correspond to significance at the 10% level, may also be reported to provide context. *What Researchers Mean by . . . Statistical Significance*, INST. FOR WORK & HEALTH, https://www.iwh.on.ca/wrmb/statistical-significance (last visited Nov. 26, 2014).

I. BACKGROUND

The presumption of validity, which was originally a common law presumption that is now codified at 35 U.S.C. § 282,[13] has been treated by the courts as providing the normative foundation for the "clear and convincing" standard of proof for invalidating a patent.[14] The Federal Circuit views the presumption of validity and the clear and convincing standard for rebutting it as "different expressions of the same thing."[15] Indeed, in most adjudicatory contexts, separating the effect of the presumption from the standard of proof is difficult because the former is analytically subsumed in the latter. For example, if an accused infringer files a motion for summary judgment on an invalidity issue, the judge's analysis in deciding the motion will focus on whether the movant has carried his burden under the clear and convincing standard. The presumption of validity, to the extent that it is part of the judge's analysis, is inherent in the evaluation of whether the movant has satisfied the applicable standard of proof because the legal effect of the presumption is limited to placing the burden of proving invalidity on the patent challenger.[16]

Where the presumption itself (separate and apart from the clear and convincing standard) is perceived to materially affect case outcomes is during jury trials. Specifically, the conventional wisdom among practitioners and judges suggests that including an instruction on the presumption of validity communicates a powerful normative message to a lay jury about the need to respect the decisions of the PTO.[17]

13. 35 U.S.C. § 282(a) (2012) ("A patent shall be presumed valid.").

14. Microsoft Corp. v. i4i Ltd. P'ship, 564 U.S. ___, ___, 131 S. Ct. 2238, 2246 (2011).

15. *Chiron*, 363 F.3d at 1258 ("[T]he presumption of validity and heightened burden of proving invalidity 'are static and in reality different expressions of the same thing—a single hurdle to be cleared.'" (quoting Am. Hoist & Derrick Co. v. Sowa & Sons, Inc., 725 F.2d 1350, 1360 (Fed. Cir. 1984))).

16. 35 U.S.C. § 282(a); New England Braiding Co., Inc. v. A.W. Chesterton Co., 970 F.2d 878, 882 (Fed. Cir. 1992) ("The presumption acts as a procedural device which places the burden of going forward with evidence and the ultimate burden of persuasion of invalidity at trial on the alleged infringer.").

17. *See, e.g.*, David C. Bohrer, *Knocking the Eagle Off the Patent Owner's Shoulder:* Chiron *Holds that Jurors Don't Have to Be Told that a Patent Is Presumed Valid*, 21 SANTA CLARA COMPUTER & HIGH TECH. L.J. 259, 282–83 (2004) ("The gospel among patent trial attorneys is that jurors are extremely reluctant to second-guess the examiner . . . .

In a patent trial, the presumption of validity is typically mentioned in the jury instructions in conjunction with the clear and convincing standard for proving invalidity.[18] Under Federal Circuit law, the presumption need *not* be explicitly mentioned to jurors so long as they are informed that the burden rests on the accused infringer to prove invalidity by clear and convincing evidence.[19] This is because the presumption, which has no evidentiary value,[20] is simply a procedural device that assigns the burden of proof [21] and is considered to be part of the same hurdle imposed by the clear and convincing standard.[22] It may appear then, that mentioning the presumption is essentially redundant if the jury instructions already recite the clear and convincing standard. However, the perceived value to the patentee of instructing the jury on the presumption appears to lie in creating an atmosphere in the courtroom that discourages jurors from second-guessing the PTO.[23] That is, the "expressive function" of the presumption of validity[24] may take a more salient role during trial, when the jury is present, than during the pretrial stage. As Mark Janis has observed, the presumption itself carries an

Why should jurors feel they cannot second-guess the examiner? The answer is clear: this is what they are told to feel by the instruction on the presumption and related arguments by counsel." (emphasis omitted)).

18. *See, e.g.*, AM. INTELL. PROP. LAW ASS'N, AIPLA'S MODEL PATENT JURY INSTRUCTIONS 9 (2012), *available at* http://www.aipla.org/learningcenter/library/books/other-pubs/Documents/2012%20final%20model%20jury%20instructions.docx.

19. *Chiron*, 363 F.3d at 1258–59 (internal citations omitted).

20. SSIH Equip. S.A. v. U.S. Int'l Trade Comm'n, 718 F.2d 365, 375 (Fed. Cir. 1983) ("The presumption of validity afforded by 35 U.S.C. § 282 does not have independent evidentiary value. Rather the presumption places the burden of going forward, as well as the burden of persuasion, upon the party asserting invalidity." (citing Solder Removal Co. v. U.S. Int'l Trade Comm'n, 582 F.2d 628, 632 (CCPA 1978))).

21. 35 U.S.C. § 282(a) ("The burden of establishing invalidity of a patent or any claim thereof shall rest on the party asserting such invalidity."); Stratoflex, Inc. v. Aeroquip Corp., 713 F.2d 1530, 1534 (Fed. Cir. 1983) ("The presumption, like all legal presumptions, is a procedural device, not substantive law. It does require the decisionmaker to employ a decisional approach that starts with acceptance of the patent claims as valid and that looks to the challenger for proof of the contrary.").

22. *Chiron*, 363 F.3d at 1258–59.

23. *See, e.g.*, Bohrer, *supra* note 17, at 282–83.

24. *See generally* Mark D. Janis, *Reforming Patent Validity Litigation: The "Dubious Preponderance,"* 19 BERKELEY TECH. L.J. 923, 927 (2004) ("[A]cknowledgment of the presumption's expressive function reminds us that the fact that we *have* a presumption of patent validity is as significant as the precise verbal formulation that we use for the standard of evidence for overcoming the presumption."). For a general discussion of the law's expressive function, see Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021 (1996).

overlying message that has significance independent of the standard of proof.[25]

To patentees and accused infringers alike, an explicit statement in the jury instructions that a patent is presumed valid is not simply a "different expression"[26] of the applicable standard of proof. Rather, they view it as a powerful mechanism for injecting pro-patentee bias, particularly because it is being delivered by the judge, who is the sole neutral authority-figure in the courtroom.[27]

While jurors are commonly perceived to be highly deferential to the PTO,[28] it is unclear to what extent that deference may be attributable to jurors feeling strongly discouraged from second-guessing the PTO upon being instructed on the presumption. In addition, it is possible that lay individuals might be confusing the presumption with evidence.[29] Such concerns have been recognized by some judges and practitioners, who have prepared alternative model patent jury instructions that do not mention the presumption.[30] Nevertheless, expressly informing the jury of the presumption of validity is common practice: the model jury instructions of at least one influential circuit,[31] as well as the instructions prepared by certain national intellectual property bar organizations such as the American Intellectual Property

25. Janis, *supra* note 24, at 930 ("[T]here is no strict, inevitable correlation between the words of the evidentiary standard and the overlying message delivered by the presumption of validity. The message is independently significant for purposes of patent policy . . . .").

26. *See supra* note 15 and accompanying text.

27. Childs, *supra* note 6, at 172.

28. *See* Kimberly A. Moore, *Juries, Patent Cases, & a Lack of Transparency*, 39 HOUS. L. REV. 779, 787 (2002) ("[P]ractitioners and scholars alike have frequently opined that juries are not likely to invalidate patents because juries favor inventors and are unlikely to second-guess the Patent Office that has technically trained examiners who already issued the patents.").

29. *See* NAT'L JURY INSTRUCTION PROJECT, MODEL PATENT JURY INSTRUCTIONS 33 (2009), *available at* http://www.nationaljuryinstructions.org/documents/NationalPatentJuryInstructions.pdf ("[I]nstructing the jury on the presumption in addition to informing it of the highly probable burden of proof may cause jury confusion as to its role in deciding invalidity.").

30. *See, e.g.*, MODEL PATENT JURY INSTRUCTIONS FOR THE NORTHERN DISTRICT OF CALIFORNIA (2014), *available at* http://www.cand.uscourts.gov/juryinstructions; NAT'L JURY INSTRUCTION PROJECT, *supra* note 29, at 33.

31. *See* FEDERAL CIVIL JURY INSTRUCTIONS OF THE SEVENTH CIRCUIT 228 (2009), *available at* https://www.ca7.uscourts.gov/Pattern_Jury_Instr/7th_civ_instruc_2009.pdf.

Law Association ("AIPLA")[32] and the Federal Circuit Bar Association ("FCBA")[33]—which reflect the prevailing "best practices" among practitioners—mention both the presumption and the clear and convincing standard.

To level the playing field against the patentee in front of the jury, some accused infringers file motions in limine to exclude any mention of the presumption,[34] while others attempt (with little success) to introduce evidence during trial on the operational realities of the PTO (e.g., patent quality issues, the application backlog, funding issues).[35] Patentees, for their part, file motions in limine to bar accused infringers from mentioning anything to the jury that may disparage the PTO.[36] Although trial judges usually exclude evidence or arguments critical of the PTO on the ground that such information would be highly prejudicial and would undermine the presumption of validity,[37] some judges have reserved the right to allow such information if the patentee tries to argue to the jury that deference to the PTO is owed at a

32. AM. INTELL. PROP. LAW ASS'N, *supra* note 18, at 9.

33. FED. CIRCUIT BAR ASS'N, MODEL PATENT JURY INSTRUCTIONS 36, 48 (2012), *available at* http://memberconnections.com/olc/filelib/LVFC/cpages/9004/Library/2012%20Updated%20FCBA%20Model%20Patent%20Jury%20Instructions.pdf.

34. *See, e.g.*, ATC's Memorandum of Points and Authorities in Support of its Motion *in Limine* No. 1 to Preclude References to the Presumption of Validity of the '356 Patent at 3, Presidio Components Inc. v. Am. Technical Ceramics Corp., No. 3:08-cv-00335-IEG-NLS (S.D. Cal. Oct. 19, 2009), ECF No. 209-1; *see also* Voda v. Cordis Corp., No. CIV-03-1512-L, 2006 U.S. Dist. LEXIS 97321, at *7 (W.D. Okla. May 10, 2006) (granting defendant's motion to preclude plaintiff from referring to the presumption of validity).

35. *See, e.g.*, Neutrino Dev. Corp. v. Sonosite, Inc., 410 F. Supp. 2d 529, 544 (S.D. Tex. 2006) ("The Court finds that, to the extent that [the defendant's expert] testimony simply addresses the potential pressures and potential for error at the PTO, such testimony is inadmissible."); TAP Pharm. Prods., Inc. v. Owl Pharms., L.L.C., No. 1:99 CV 2715, 2003 U.S. Dist. LEXIS 27657, at *5 (N.D. Ohio Feb. 18, 2003) ("Testimony by [the defendant's expert] about the relative shortage of patent examiners at the PTO is inadmissible. The only purpose such testimony would serve would be to undermine the presumption of validity of the patents-in-suit." (internal citation omitted)).

36. *See, e.g.*, Cook's Motion *in Limine* No. 5: To Preclude Endologix from Offering Argument or Evidence About the Competence of the Examination Process in the U.S. Patent and Trademark Office, or from Otherwise Denigrating the Office, Its Examiners, or the Examination Process at 2, Cook Inc. v. Endologix, Inc., No. 1:09-cv-1248-TWP-DKL (S.D. Ind. Aug. 10, 2012), ECF No. 255.

37. *See, e.g.*, Bausch & Lomb, Inc. v. Alcon Labs., Inc., 79 F. Supp. 2d 252, 255 (W.D.N.Y. 2000) ("I find such testimony [concerning problems at the PTO] to be inadmissible. It appears that the purpose of this testimony would be to attempt to undermine the presumption of validity under 35 U.S.C. § 282 by inviting the jury to speculate about possible defects, errors, or omissions in the application process . . ."); Applied Materials, Inc. v. Advanced Semiconductor Materials Am., Inc., No. C 92-20643 RMW, 1995 WL 261407, at *3 (N.D. Cal. Apr. 25, 1995).

level beyond that required by the law.[38] The Federal Circuit generally views arguments criticizing the PTO to be improper in front of a jury, and will order a new trial if warranted by the totality of the circumstances.[39]

Whether and under what circumstances the presumption of validity (or, in some cases, criticisms of the PTO) should be mentioned to the jury is a question for which empirical analysis may be useful in helping to test the conventional assumptions that drive trial strategy. Despite this need, there appears to be only one prior empirical study that has attempted to collect data on the presumption of validity separately from the clear and convincing standard.[40] However, that study did not focus on jury trials, but rather on Federal Circuit decisions.[41] In addition, it used a sample size that was too small to allow any potential impact of the presumption to be reliably assessed separately from that of the evidentiary standard of proof.[42]

Although the impact of a presumption instruction has not been specifically analyzed in previous empirical research relating to jury trials, the clear and convincing evidentiary standard has been the subject of a recent experiment with mock jurors. In 2013, David Schwartz and Christopher Seaman published a

38. *See, e.g.*, *Bausch & Lomb*, 79 F. Supp. 2d at 256 ("I caution plaintiff that if it opens the door by suggesting that some extraordinary deference is due in this case, the court may revisit this ruling [barring argument concerning the PTO's problems]." (internal quotations and citation omitted)); *Applied Materials*, 1995 WL 261407, at *3 ("The court will reconsider this ruling [barring mention of the PTO's operational realities], however, if Applied opens the door by presenting evidence suggesting that some extraordinary deference is due in this case.").

39. *See, e.g.*, Novo Nordisk A/S v. Becton Dickinson & Co., 304 F.3d 1216, 1220 (Fed. Cir. 2002) ("[O]n balance, we conclude that a new trial is not warranted in the circumstances that here prevailed, for the issue[] of examiner competence . . . [was] not raised by post-trial motion; this inaction . . . suggests that in the overall context of the two-week trial, these aspects were less inflammatory than [the patentee] now maintains.").

40. *See* Etan S. Chatlynne, *UPDATE: Investigating Patent Law's Presumption of Validity*, PATENTLY-O (Nov. 7, 2010), http://patentlyo.com/patent/2010/11/update-investigating-patent-laws-presumption-of-validity.html [hereinafter "Chatlynne Update"] (This blog post is an update of the results reported in Etan S. Chatlynne, *Investigating Patent Law's Presumption of Validity—An Empirical Analysis*, 2010 PATENTLY-O PAT. L.J. 37 (2010), *available at* http://patentlyo.com/media/docs/2010/03/chatlynne.presumptionofvalidity.final.pdf).

41. *See* Chatlynne Update, *supra* note 40.

42. In a dataset compiling 119 invalidity challenges, Chatlynne reported that the Federal Circuit expressly applied the presumption of validity or the evidentiary standard in its analysis a total of twenty-six times. *Id*.

study in which they used a survey experiment to investigate the effect of modifying the standard of proof on a juror's decision to find a patent invalid.[43] They presented mock jurors with a patent case hypothetical where the ultimate issue to be decided was obviousness.[44] The mock jurors were then randomly assigned to one of three jury instructions that contained different versions of the standard of proof: (1) clear and convincing evidence; (2) clear and convincing evidence with an additional instruction based on *Microsoft Corp. v. i4i Ltd. Partnership*[45] pertaining to new evidence not considered by the PTO; and (3) preponderance of the evidence.[46] The results of the experiment by Schwartz and Seaman suggest that jurors' decisions to find invalidity may be affected substantially by the standard of proof.[47]

Given that jury instructions on the clear and convincing standard have been the subject of experimental analysis, a logical next step would be to explore the effect of instructing the jury on the presumption of validity.

II. RESEARCH DESIGN

The conventional assumptions pertaining to the inclusion of an instruction on the presumption of validity in the jury instructions raise a variety of normative questions, some of which may be amenable to experimental study. This article seeks to explore two such questions. *First*, would mentioning the presumption to the jury actually affect the likelihood that a patent will be found invalid? (If so, what is the magnitude of that impact, given that juries are perceived to be generally deferential to the PTO and view inventors positively?)[48] *Second*, if the accused infringer were to introduce information critical of the PTO during trial, what, if any, impact could this have on the presumption of validity?

43. David L. Schwartz & Christopher B. Seaman, *Standards of Proof in Civil Litigation: An Experiment from Patent Law*, 26 HARV. J.L. & TECH. 429, 431–32 (2013).

44. *Id.* at 451.

45. 564 U.S. ___, ___, 131 S. Ct. 2238 (2011).

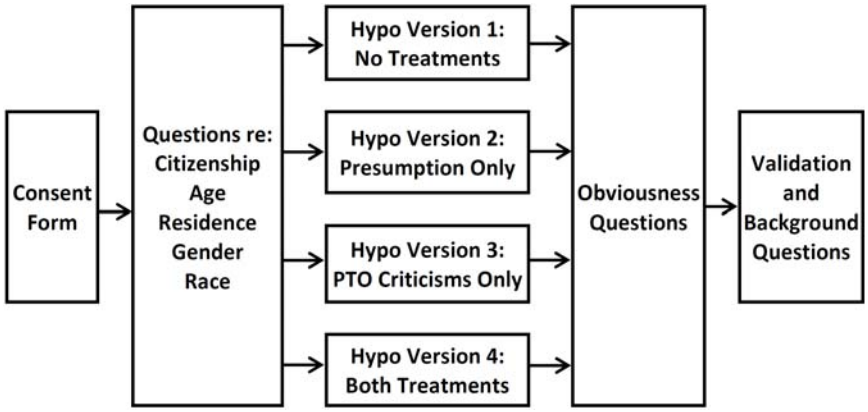
46. Schwartz & Seaman, *supra* note 43, at 432.

47. *Id*. at 459–61.

48. Moore, *supra* note 28, at 787.

To explore these issues, an online survey experiment[49] was conducted in which mock jurors were presented with a hypothetical patent case and were asked whether the asserted patent was invalid for obviousness. To mitigate potential response bias and "demand effects,"[50] a "between-subjects" design[51] was used for the survey: Each respondent was allowed to take the survey only once, and there were no questions that asked about the same issue both before and after the hypothetical. The general flow of the survey experiment is shown below:

Figure 1: Flow Diagram of Survey Experiment



As shown in Figure 1, the survey experiment begins with the consent form and elicits basic demographic information about the respondent who will serve as a mock juror.[52] The respondent is

49. A survey experiment is different from a regular survey in that it involves a "treatment" component, which is an element of the survey that is systematically varied in relation to a "control" or a baseline, so as to allow causal inferences to be drawn. *See* Shari Seidman Diamond, *Reference Guide on Survey Research*, *in* FED. JUDICIAL CTR., REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 359, 397–98, 421 (3d ed. 2011).

50. *See* Rachel Croson, *Why and How to Experiment: Methodologies from Experimental Economics*, 2002 U. ILL. L. REV. 921, 933 ("[T]he researcher must be careful to avoid *demand effects*—avoid suggesting the desired results to the subjects either explicitly or implicitly.").

51. *See* Christopher Slobogin & Lauren Brinkley-Rubinstein, *Putting Desert in Its Place*, 65 STAN. L. REV. 77, 100 n.101 (2013) ("In a between-subjects design (to be distinguished from a 'within-subjects' design), the manipulation is hidden from the subjects; its effect is studied by using two or more samples, ideally matched in all relevant respects, with each sample receiving a different independent variable . . . .").

52. The presentation of a few basic demographic questions at the beginning of the

then randomly assigned to one of four versions of the hypothetical.[53] Each version of the hypothetical presents the identical fact pattern except for the selective presence (or absence) of either the presumption of validity in the jury instructions or criticisms about the PTO in the accused infringer's arguments, or both. Afterwards, the respondent is asked whether the patent described in the hypothetical is invalid for obviousness.[54] The respondent is then presented with validation questions that test whether he understood basic facts about the hypothetical—if the respondent answers the validation questions incorrectly, his answers would be excluded from the analysis. Finally, the survey concludes with questions that ask about the respondent's background, such as patent-related experience, education, and jury service.

The hypothetical and the associated questions relating to the obviousness issue were adapted from Schwartz and Seaman's "standards of proof" experiment.[55] The hypothetical, which describes a patent dispute over a golf ball design, was presented in three parts: the overview, the parties' arguments, and the jury instructions for deciding whether the asserted patent was invalid[56] for obviousness. Using Schwartz and Seaman's hypothetical provided several advantages. *First*, it was already field-tested as being reasonably understandable to lay subjects

survey experiment prior to the hypothetical is intended to track the general order of events in a trial setting, in which prospective jurors are asked various demographic and background questions during the jury selection process. This ordering has been used in published studies of experiments with mock jurors. *See, e.g.*, Shari Seidman Diamond, Michael J. Saks & Stephan Landsman, *Juror Judgments About Liability and Damages: Sources of Variability and Ways to Increase Consistency*, 48 DEPAUL L. REV. 301, 304–05 (1998) ("Before viewing a videotape of the trial . . . , each juror filled out a questionnaire providing the kind of information that jurors might be asked to provide during jury selection in a case of this type."). For a sample juror questionnaire, see JUROR QUESTIONNAIRE, FED. JUDICIAL CTR. 4–17, http://www.fjc.gov/public/pdf.nsf/lookup/dpen0023.pdf/$file/dpen0023.pdf (last visited Nov. 26, 2014). In this study, in order to avoid respondent exhaustion prior to reviewing the hypothetical, only five basic demographic questions (citizenship, age, residence, gender, race) were asked prior to the hypothetical, with additional demographic/background questions presented at the end of the survey.

53. See *infra* Appendix 1 for an annotated version of the hypothetical.

54. *See infra* Appendix 2.

55. Schwartz & Seaman, *supra* note 43, at 451–56, 474–78.

56. In general, invalidity is evaluated on a claim-by-claim basis. *See* 35 U.S.C. § 282(a) (2012). For simplicity, however, the hypothetical did not identify any specific claims such that the obviousness questions asked whether the patent (as opposed to a claim) was invalid. *See infra* Appendix 2.

without any background in science or engineering.[57] *Second*, the hypothetical was based on a case that had two jury verdicts that reached opposite conclusions concerning invalidity, which may indicate that there is no clear "right" answer.[58] *Finally*, the similarities between the hypothetical used in this experiment, which explores the presumption of validity, with that used in Schwartz and Seaman's experiment, which explores the clear and convincing standard,[59] may facilitate comparisons between the two studies.

For the purposes of this study, Schwartz and Seaman's hypothetical was modified as follows:

- A single standard of proof (clear and convincing) was recited in the instructions to the mock jurors for deciding the invalidity issue, as opposed to the three different standards used in Schwartz and Seaman's study.[60]

- Two treatment blocks were added in order to test the effects of mentioning the presumption of validity and criticisms of the PTO to the jury.

- In Schwartz and Seaman's study, the hypothetical specified that a key prior art reference was *not* considered by the examiner, which allowed them to test different versions of the standard of proof.[61] In the present study, the hypothetical was modified to state that the prior art reference *had been* considered, in order to better gauge the mock jurors' baseline level of deference to the PTO, assuming that the PTO had not made any glaring errors or omissions.

The two treatment blocks used in the hypothetical were: (i) an explanation of the presumption of validity in the section providing instructions to the mock jurors for deciding the invalidity issue; and (ii) a description of common criticisms of the PTO in the accused infringer's argument section.[62] The treatment blocks are reproduced below:

57. Schwartz & Seaman, *supra* note 43, at 451.

58. *Id*.

59. *See supra* notes 43–47 and accompanying text.

60. Schwartz & Seaman, *supra* note 43, at 432.

61. *Id*. at 453.

62. If a court were to ever allow information critical of the PTO to be presented, it would most likely be presented by a patent law expert called by the accused infringer.

Presumption Treatment Block:[63]

> *Under the law, Acme's patent is presumed to be valid. In other words, it is presumed to have been properly granted. When a party attacking the validity of a patent relies on prior art that was specifically considered by the patent examiner, that party bears the burden of overcoming the deference due a qualified government agency official who is presumed to have performed his or her job correctly. The presumption of validity that is accorded a duly-issued patent can be overcome by "clear and convincing" evidence of obviousness.*

PTO Criticisms Treatment Block:[64]

> *That the patent examiner might have made a mistake should not be surprising. As recognized by numerous academic researchers, poor patent quality is a serious problem. The PTO is underfunded and has a backlog of approximately 600,000 patent applications that are awaiting examination.*[65] *The patent examiners are overworked, and are simply not given enough time to review patent applications thoroughly.*[66] *Indeed, according to one academic study, about half of all patents that are litigated in court are found to be invalid.*[67]

The selective inclusion of the treatment blocks yielded four versions of the hypothetical to which the mock jurors were randomly assigned:

63. The presumption of validity treatment block is an amalgam of the relevant language from the AIPLA and FCBA model jury instructions. *See supra* notes 32–33.

64. The footnotes to supporting sources were not included in the versions of the hypothetical presented to the mock jurors.

65. In fiscal year 2013, 616,409 patent applications were awaiting a first action by an examiner. U.S. PATENT AND TRADEMARK OFFICE, U.S. PATENT AND TRADEMARK OFFICE PERFORMANCE & ACCOUNTABILITY REPORT 191 (2013), *available at* http://www.uspto.gov/about/stratplan/ar/USPTOFY2013PAR.pdf.

66. FED. TRADE COMM'N, To PROMOTE INNOVATION: THE PROPER BALANCE OF COMPETITION AND PATENT LAW AND POLICY 8–10 (2003) [hereinafter "FTC Report"], *available at* http://www.ftc.gov/sites/default/files/documents/reports/promote-innovation-proper-balance-competition-and-patent-law-and-policy/innovationrpt.pdf.

67. John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 205–06 (1998).

(1) *No Treatments* version: Neither the presumption of validity nor criticisms of the PTO were included.

(2) *Presumption Only* version: The presumption was included but PTO criticisms were not.

(3) *PTO Criticisms Only* version: PTO criticisms were included, but the presumption was not.

(4) *Both Treatments* version: Both the presumption and PTO criticisms were included.

An annotated version of the hypothetical showing the treatment blocks is provided in Appendix 1. Whereas the presumption of validity was contained in a treatment block, all versions of the hypothetical mentioned the clear and convincing standard of proof, as required under Federal Circuit law.[68]

The mock jurors were recruited via Amazon Mechanical Turk ("MTURK"), a website run by Amazon.com, Inc., where individuals may sign up to perform online "Human Intelligence Tasks" for pay.[69] MTURK is a popular platform for social science survey research.[70] The respondent sample was limited to the demographic profile of jury-eligible adults: United States citizens who are at least eighteen years old, and who are currently residing in the United States.[71] The MTURK site readily allows the respondent pool to be restricted to individuals who are at least eighteen years of age who reside in the United States because anyone who signs up to work on MTURK must provide verification of his or her age and residency.[72] Data on other

68. *See* Chiron Corp. v. Genentech, Inc., 363 F.3d 1247, 1258–59 (Fed. Cir. 2004).

69. AMAZON MECH. TURK, https://www.mturk.com/mturk/welcome (last visited Nov. 26, 2014).

70. Several recent experimental studies in the intellectual property field have used MTURK. *E.g.*, Christopher Buccafusco, Zachary C. Burns, Jeanne C. Fromer & Christopher Jon Sprigman, *Experimental Tests of Intellectual Property Laws' Creativity Thresholds*, 92 TEX. L. REV. 1921, 1950 (2014); Christopher Buccafusco & Paul J. Heald, *Do Bad Things Happen When Works Enter the Public Domain? Empirical Tests of Copyright Term Extension*, 28 BERKELEY TECH. L.J. 1, 24 (2013); Christopher Jon Sprigman, Christopher Buccafusco & Zachary Burns, *What's a Name Worth? Experimental Tests of the Value of Attribution in Intellectual Property*, 93 B.U. L. REV. 1389, 1405 (2013); Schwartz & Seaman, *supra* note 43, at 456.

71. *Juror Qualifications, Exemptions and Excuses*, U.S. COURTS, http://www.uscourts.gov/FederalCourts/JuryService/JurorQualificaitons.aspx (last visited Nov. 26, 2014).

72. Each worker who registers on MTURK must provide verification of their residence in order for MTURK to process tax information. *See Amazon Mechanical Turk*

demographic characteristics, including citizenship, were obtained through self-identification. Although MTURK allows the respondent pool to be further restricted based on a respondent's general approval rating on the site, this option was not used in order to allow a wide cross-section of respondents to participate in the study. Each respondent was paid $1.00 for successfully completing the survey experiment, which ran on MTURK for two days in July 2014.

Initially, 2616 respondents accessed at least the first page of the survey, of which 2412 jury-eligible respondents progressed through survey termination. Of these respondents, 667 were eliminated because of quality issues that would render their responses unreliable, such as speeding through the survey,[73] failing to correctly answer basic factual questions about the hypothetical, and providing logically inconsistent answers to certain questions.[74] This yielded 1745 respondents for analysis. The respondent tally for each of the four treatment versions of the hypothetical is shown below:

Participation Agreement, AMAZON MECH. TURK (Nov. 1, 2012), https://www.mturk.com/mturk/conditionsofuse.

73. Because the hypothetical related to a topic (patent law) that may be unfamiliar to most adults, it was important that the respondents read at a pace that allowed for comprehension. According to one measure, the average adult reads at the rate of 300 words per minute, while the average college professor reads at the rate of 675 words per minute. Brett Nelson, *Do You Read Fast Enough to Be Successful?*, FORBES (June 4, 2012, 9:09 AM), http://www.forbes.com/sites/brettnelson/2012/06/04/do-you-read-fast-enough-to-be-successful/. Respondents who read each page of the patent case hypothetical faster than three times the average adult (i.e., 900 words per minute) had their responses eliminated from the analysis.

74. For example, if a respondent specified in one question that he did not serve on a jury but specified in another question that he served as a juror in a civil case, his answers were eliminated from the final analysis.

Table 1

Treatment Scenario	Respondents (Mock Jurors)
No Treatments	441
Presumption Only	430
PTO Criticisms Only	436
Both Treatments	438
Total	**1,745**

The mean age of the respondents was thirty-four years, while the median age was thirty. They were 49% female and 78% white. A majority (59%) had at least a college degree.[75] Recent demographic statistics of federal juries are not available for comparison.[76] However, a few statistics from a study conducted in 2004[77] of people who reported for jury duty in King County, Washington (which has a population of over two million),[78] may be instructive. According to that study, the individuals who appeared for jury duty in county court (N=1545)[79] had a median age of forty-eight, 69% were college graduates, 54% were female, and 86% were white.[80] In contrast, the general county census indicated that its residents had a median age of forty-six, 43% were college graduates, 51% were female, and 74% were white.[81] Notably, this 2004 study revealed that a substantially higher percentage of individuals who showed up for jury duty had college degrees compared to the general population (69% versus 43%).[82] In addition, the demographics of those who were eventually sworn in as jurors were similar to those who showed up for jury duty.[83] When the King County study is compared to the MTURK

75. This tally does not include individuals who attended college without obtaining a degree.

76. *See* Schwartz & Seaman, *supra* note 43, at 458 n.183.

77. JOHN GASTIL ET AL., THE JURY AND DEMOCRACY: HOW DELIBERATION PROMOTES CIVIC ENGAGEMENT AND POLITICAL PARTICIPATION 53 (2010).

78. *State and County QuickFacts: King County, Washington*, U.S. CENSUS BUREAU, http://quickfacts.census.gov/qfd/states/53/53033.html (last visited Nov. 26, 2014) (listing King County's population estimate as of 2013).

79. GASTIL ET AL., *supra* note 77, at 65 tbl.4.3.

80. *Id.* at 61 tbl.4.2.

81. *Id*.

82. *Id*.

83. *Id.*

respondents, the most salient difference is the median age, whereby the MTURK respondents are substantially younger as a group. With an age distribution that is heavily skewed toward younger respondents, caution is warranted when analyzing the relationship between case outcomes and the age variable[84] as well as other background variables (e.g., jury service)[85] that have some degree of age-dependence.

III. RESULTS

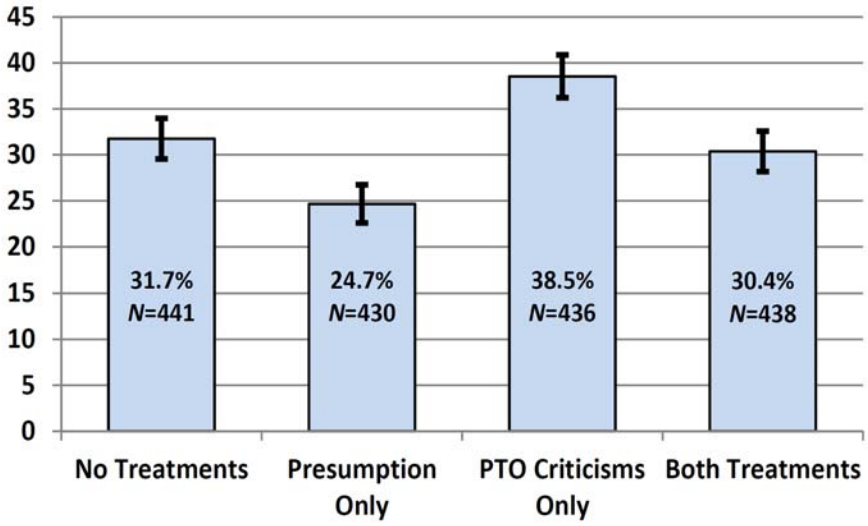
Overall, the results largely confirm the conventional assumptions held by litigants on the likely effect of informing the jury about the presumption of validity and criticisms about the PTO. At a high-level, the data reveal that mock jurors who were exposed to the presumption instruction were significantly less likely to find invalidity. Conversely, informing them of various criticisms of the PTO appeared to have an effect of comparable magnitude in the opposite direction. A notable but unexpected result is that when both the presumption and PTO criticisms were presented, their effects seemingly canceled each other. A detailed presentation of the data follows.

A. *Treatment Scenarios*

The rates at which the mock jurors found invalidity based on obviousness are graphically summarized below for each of the four treatment scenarios:

84. *See infra* notes 116–17 and accompanying text.
85. *See infra* notes 128–31 and accompanying text.

Figure 2: Invalidity Decisions
(Percentages with Standard Error Bars)



As an initial step, each pair of treatment scenarios was compared using a Chi-square test to determine if the differences in the invalidity rates were statistically significant.[86] A comparison of the *No Treatments* and the *Presumption Only* scenarios (both of which do not contain any criticisms about the PTO) shows a statistically significant drop in invalidity decisions when the presumption is mentioned (31.7% vs. 24.7%; p=0.020).[87] If criticisms about the PTO were present in both of the scenarios being compared, a statistically significant drop in invalidity decisions occurred if the presumption was added, as shown by a comparison of the *PTO Criticisms Only* scenario and the *Both Treatments* scenario (38.5% vs. 30.4%; p=0.011).[88] These results lend support to the belief of accused infringers that instructing the jury on the presumption of validity may decrease the likelihood of an invalidity finding.

As for the conventional assumption that criticisms about the PTO are highly prejudicial, the results appear to confirm this.

86. *See generally* GRAHAM CURRELL & ANTONY DOWMAN, ESSENTIAL MATHEMATICS AND STATISTICS FOR SCIENCE 245 (2005).

87. *See supra* Figure 2.

88. *Id*.

The difference in the invalidity rates between the *No Treatments* and the *PTO Criticisms Only* scenarios is statistically significant (31.7% vs. 38.5%; p=0.035).[89] In addition, a comparison of the typical patent case scenario where the presumption instruction is given (*Presumption Only*) with the scenario where, in addition to the presumption, PTO criticisms are also introduced (*Both Treatments*) reveals a difference in the invalidity rate that is significant at the 10% level (24.7% vs. 30.4%; p=0.060).[90]

Of the pair-wise comparisons, perhaps the most intriguing result is the comparison of the *No Treatments* and the *Both Treatments* scenarios, where the former has neither the presumption instruction nor the PTO criticisms, while the latter has both treatments. It appears as if the effects of the two treatments cancel each other (31.7% vs. 30.4%; p=0.658).[91] This result was unexpected, given that negative information is generally deemed to carry more weight and exert a stronger influence than either positive or neutral information.[92]

The difference between the *Presumption Only* and the *PTO Criticisms Only* scenarios was highly statistically significant (24.7% vs. 38.5%; $p < 0.001$).[93] This result was not unexpected, given that a comparison of these two scenarios does not reflect the incremental presence (or absence) of a single treatment, but rather a direct comparison of the effects of two different treatments.

To confirm whether the significance levels reported by the pair-wise Chi-square comparisons would continue to hold after controlling for various demographic and background characteristics of the mock jurors, a series of multiple logistic regression models were used, as shown in Appendices 3 through 6.[94] The dependent variable corresponds to a finding of invalidity

89. *Id*.

90. *Id*.

91. *Id*.

92. *See* Roy F. Baumeister, Ellen Bratslavsky, Catrin Finkenauer, Kathleen D. Vohs, *Bad Is Stronger Than Good*, 5 REV. GEN. PSYCHOL. 323, 323 (2001) ("When equal measures of good and bad are present, however, the psychological effects of bad ones outweigh those of the good ones.").

93. *See supra* Figure 2.

94. *See generally* Chao-Ying Joanne Peng, Kuk Lida Lee & Gary M. Ingersoll, *An Introduction to Logistic Regression Analysis and Reporting*, 96 J. EDUC. RES. 3, 4 (2002) ("Generally, logistic regression is well suited for describing and testing hypotheses about

by reason of obviousness. The predictor variables were the different treatment scenarios and various personal characteristics of the respondents. Overall, the regression models confirmed the results of the pair-wise comparisons.

In the regression models, each of the four treatment versions of the hypothetical was represented by a dummy variable, where one of the versions served as the base variable to which the other three were compared. For this reason, four general models were used for the logistic regression:

Table 2: Models and Base Comparison Variables

Model[95]	Base Comparison Variable
A	No Treatments
B	Presumption Only
C	PTO Criticisms Only
D	Both Treatments

Of the four Models, Models A and B are of particular interest. In Model A, the *No Treatments* scenario is the base comparison variable to which the other treatment scenarios are compared. Because the *No Treatments* scenario does not have any treatments, it may serve as an intuitive baseline against which the effects of the other treatments may be evaluated. Model B is also notable because the base comparison variable (i.e., the *Presumption Only* scenario) corresponds to a common scenario in actual patent trials in which the jury is informed of the presumption but is not provided any information critical of the PTO.

Each Model consists of four sub-models, numbered 1 through 4. Sub-models A1–A4, B1–B4, C1–C4, and D1–D4 use the same corresponding sets of predictor variables in the regression, except for the base comparison variable for the treatment scenarios. The

relationships between a categorical outcome variable and one or more categorical or continuous predictor variables.").

95. For the corresponding logistic regression tables, see *infra* Appendices 3–6.

sub-models were created to mitigate or avoid potential multicollinearity issues.[96] Specifically, the variable "College Graduate," which indicates whether a respondent's level of education is at least a college degree, is in a separate sub-model from one that uses the variable "Science Degree," which indicates that the respondent has a college or graduate degree in science, engineering, or mathematics. Similarly, the variable "Jury Service," which indicates whether a respondent has served on a jury, is in a different sub-model from the variables "Civil Juror" and "Criminal Juror,"[97] which indicate whether a respondent served on a civil jury or on a criminal jury, respectively. In total, four sub-models were used to capture the various combinations of alternative variables (i.e., "College Graduate" vs. "Science Degree"; "Jury Service" vs. "Civil Juror" and "Criminal Juror").

Turning now to the results of the multiple logistic regression, Model A reveals that, when compared to the *No Treatments* scenario, the *Presumption Only* scenario decreased the odds, by a statistically significant margin, that the mock juror in this study would find invalidity.[98] By contrast, the *PTO Criticisms Only* scenario *increased* the odds of an invalidity finding by a statistically significant margin.[99] The *Both Treatments* scenario did not result in a statistically significant change in the odds.[100]

96. *See* ANDREW SIEGEL, PRACTICAL BUSINESS STATISTICS 372 (6th ed. 2012) (explaining that multicollinearity makes it "difficult for multiple regression to distinguish between the effect of one variable and the effect of another").

97. Twenty-one respondents served in both civil and criminal cases—they were counted in both the "Civil Juror" and the "Criminal Juror" variables. Fifteen respondents served on a jury but were not sure of the type of case—they were included in the "Jury Service" variable, but not in the "Civil Juror" or the "Criminal Juror" variables.

98. Across Models A1–A4, the odds ratio associated with the *Presumption Only* scenario ranged between .6965 and .7022, and the p-value ranged between 0.018 and 0.020. *See infra* Appendix 3. An odds ratio that is greater than 1.0 refers to an increase in the odds, while an odds ratio that is less than 1.0 refers to a decrease in the odds. *An Introduction to Odds, Odds Ratios and Exponents*, RESTORE @ NATIONAL CENTRE FOR RESEARCH METHODS (Jul. 25, 2011), http://www.restore.ac.uk/srme/www/fac/soc/wie/research-new/srme/modules/mod4/2/.

99. Across Models A1–A4, the odds ratio associated with the *PTO Criticisms Only* scenario ranged between 1.3604 and 1.3717, and the *p*-value ranged between 0.027 and 0.031. *See infra* Appendix 3.

100. Across Models A1–A4, the odds ratio associated with the *Both Treatments* scenario ranged between .9329 and .9376, and the *p*-value ranged between 0.636 and 0.660. *See infra* Appendix 3.

These relationships confirm the pair-wise Chi-square analysis reported earlier.[101]

The results for Model B tell a similar story. Compared to the *Presumption Only* scenario, which served as the base variable, the *No Treatments* scenario resulted in a statistically significant increase in the odds of an invalidity finding.[102] Unsurprisingly, the *PTO Criticisms Only* scenario increased the odds of an invalidity finding by a highly statistically significant margin when compared to the *Presumption Only* scenario.[103] The *Both Treatments* scenario also increased the odds of an invalidity finding, but the change was significant only at the 10% level.[104] These relationships again confirm the earlier pair-wise Chi-square comparisons.[105] The results for Models C and D similarly confirm the relationships reported by the pair-wise Chi-square analysis, albeit with a change in the significance level of the difference between the *PTO Criticisms Only* and the *Both Treatments* scenarios from the 5% level (Chi-square) to the 1% level (logistic regression).[106]

In addition to being asked to decide whether the patent in the hypothetical was invalid for obviousness, the mock jurors were also asked to specify the likelihood of obviousness on a scale of 0% (Certainly Not Obvious) to 100% (Certainly Obvious).[107] The "likelihood of obviousness" estimate provides an indication of the mock jurors' subjective impressions of the strength of the obviousness arguments presented in the hypothetical.[108] The

101. *See supra* text accompanying notes 86–87, 89, 91.

102. Across Models B1–B4, the odds ratio associated with the *No Treatments* scenario ranged between 1.4241 and 1.4357, and the *p*-value ranged between 0.018 and 0.020. *See infra* Appendix 4.

103. Across Models B1–B4, the odds ratio associated with the *PTO Criticisms Only* scenario ranged between 1.9374 and 1.9694, and $p < 0.001$. *See infra* Appendix 4.

104. Across Models B1–B4, the odds ratio associated with the *Both Treatments* scenario ranged between 1.3329 and 1.3417, and the *p*-value ranged between 0.056 and 0.061. *See infra* Appendix 4.

105. *See supra* text accompanying notes 86–87, 90, 93.

106. *See infra* Appendices 5–6; *see also supra* text accompanying note 88 (reporting the result of the Chi-Square comparison).

107. *See infra* Appendix 2.

108. The likelihood estimate also served as another way to check whether the jurors understood the hypothetical, especially the applicable standard of proof (clear and convincing). A respondent was eliminated if he answered that the patent was obvious but separately indicated that the likelihood of obviousness was less than 40%, or, alternatively, if he found nonobviousness, but indicated that the likelihood of obviousness

mean likelihood estimates are listed below, tabulated according to the treatment scenario and the disposition ("Obvious" or "Not Obvious"):

Table 3: Likelihood of Obviousness (%)

	Obvious		Not Obvious	
Treatment Scenario	***N***	**Mean (SD)**	***N***	**Mean (SD)**
No Treatments	140	**78.9** (11.1)	301	**28.5** (17.5)
Presumption Only	106	**78.7** (11.1)	324	**30.5** (17.8)
PTO Criticisms Only	168	**79.0** (10.8)	268	**30.7** (17.1)
Both Treatments	133	**80.8** (12.0)	305	**31.6** (18.5)
Welch ANOVA (Across Treatments)	***p*=0.457**		***p*=0.193**	

To see whether the treatment scenarios might affect the mock jurors' likelihood estimates, a one-way Welch ANOVA was used to compare the mean estimates across the four treatment scenarios for each disposition.[109] Notably, a statistically significant relationship could *not* be discerned between the mean likelihood estimates and the treatment scenarios, regardless of whether obviousness was found (range of means: 78.7% to 80.8%; p=0.457) or not (range of means: 28.5% to 31.6%; p=0.193).[110] This result was somewhat unexpected, given the statistically

was at least 90%. Only eighteen respondents were excluded on this basis. This question was adapted from a similar question used by Schwartz and Seaman, who used it as a check on the respondents' answers in their study. Schwartz & Seaman, *supra* note 43, at 461–62.

109. *See generally* CURRELL & DOWMAN, *supra* note 86, at 292.

110. *See supra* Table 3.

significant differences in the invalidity rates for various pairs of treatment scenarios.[111]

The final obviousness-related question asked the mock jurors to specify the level of confidence in their answers for the two prior obviousness-related questions, on a scale of one (Not Confident At All) to seven (Extremely Confident).[112] The mean levels of confidence are reported below:

Table 4: Level of Confidence in Answers

	Obvious		**Not Obvious**	
Treatment Scenario	***N***	**Mean (SD)**	***N***	**Mean (SD)**
No Treatments	140	**5.51** (1.12)	301	**5.28** (1.18)
Presumption Only	106	**5.51** (1.11)	324	**5.33** (1.17)
PTO Criticisms Only	168	**5.54** (1.06)	268	**5.29** (1.12)
Both Treatments	133	**5.40** (1.38)	305	**5.28** (1.22)
Welch ANOVA (Across Treatments)	***p*=0.797**		***p*=0.931**	

To determine whether being informed of the presumption instruction or criticisms about the PTO affected the mock jurors' overall level of confidence regarding their answers to the obviousness-related questions, a one-way Welch ANOVA was used to compare the mean confidence levels across the four treatment scenarios for each disposition. A statistically significant relationship could not be discerned between the confidence levels and the treatment scenarios, whether obviousness was found (range of means: 5.40 to 5.54; p=0.797) or not (range of means: 5.28 to 5.33; p=0.931).[113] This result was

111. *See supra* Figure 2 and text accompanying notes 86–93.
112. *See infra* Appendix 2.
113. *See supra* Table 4.

unexpected, particularly the immaterial difference in mean confidence levels between the *Presumption Only* and the *PTO Criticisms Only* scenarios.[114] In effect, the treatments did not appear to make the mock jurors feel more (or less) confident about their decisions.

B. *Other Predictors*

In addition to the treatment scenarios, the regression models included several variables based on the mock jurors' demographic and background characteristics.[115] Data on these characteristics were collected solely based on self-identification.

The regressions revealed no statistically significant effect on the odds of an invalidity decision based on age.[116] Because the median age of the respondents was thirty, it is possible that the relatively low concentration of older respondents in the sample might have prevented a statistically significant effect from being discerned.[117] Regarding gender, women were far less likely than men to find invalidity—this difference was highly statistically significant.[118] This result confirms a similar finding in Schwartz and Seaman's experiment relating to the standard of proof.[119] One potential explanation for the significance of gender might be the subject matter of the patent in the hypothetical: it is possible that men may feel more comfortable than women in second-guessing a patent examiner on an invention relating to golf. With respect to race, minorities were more likely to find invalidity than whites by a margin that was significant at the 10% level.[120] Although Schwartz and Seaman's study did not report any significance based on race,[121] the finding of significance at the 10% level in the

114. *See supra* Table 4.

115. *See infra* Appendices 3–6.

116. Across the various models, the *p*-value ranged between 0.158 and 0.248. *See infra* Appendices 3–6.

117. For example, only 146 out of 1745 respondents (8.4%) were aged fifty-five and over. In contrast, 25.8% of the U.S. population was aged fifty-five and over in 2012. U.S. CENSUS BUREAU, AGE AND SEX COMPOSITION IN THE UNITED STATES: 2012 tbl.1 (2013), *available at* http://www.census.gov/population/age/data/files/2012/2012gender_table1.xlsx.

118. Across the various models, p=0.002. *See infra* Appendices 3–6.

119. Schwartz & Seaman, *supra* note 43, at 479–80.

120. Across the various models, the *p*-value ranged between 0.053 and 0.071. *See infra* Appendices 3–6.

121. Schwartz & Seaman, *supra* note 43, at 479–80.

present study might be partly attributable to the much larger sample size (*N*=1745) compared to that of Schwartz and Seaman's study (*N*=500).[122] More generally, given the nature of the hypothetical, it is possible that the gender and race effects[123] in this study may reflect differences among the socio-demographic groups regarding their willingness to second-guess government agencies, as well as their attitudes toward golf, corporate defendants, and the patent system. A detailed exploration of the socio-demographic aspects of jury decision-making in patent cases is left to future research.

Concerning experience relevant to the subject matter of the hypothetical, the respondents were asked whether they had played golf: 955 out of 1745 (55%) had. Golf experience, however, did not have a statistically significant effect on invalidity

122. *Id*. at 456.

123. At a conference at which this article was presented, a commenter asked whether eliciting gender and race information before the hypothetical could generate "stereotype threat" that might affect the results. As mentioned previously, very basic demographic questions were posed at the beginning of the survey, similar to those found in the questionnaires that are given to prospective jurors in preparation for voir dire. *See supra* note 52. Stereotype threat is theorized to occur when an individual is put in a situation that triggers anxiety about confirming negative stereotypes about his or her socio-demographic group, such as situations that test the math ability of female students or the academic performance of black students. *See* Claude M. Steele, *A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance*, 52 AM. PSYCHOLOGIST 613, 614 (1997). However, experimental studies of the effect of inquiring about gender and ethnicity prior to taking tests have yielded mixed results. *Compare* Claude M. Steele & Joshua Aronson, *Stereotype Threat and the Intellectual Test Performance of African Americans*, 69 J. PERSONALITY & SOC. PSYCHOL. 797, 807–08 (1995) (finding stereotype threat effect), *with* Lawrence J. Stricker & William C. Ward, *Stereotype Threat, Inquiring About Test Takers' Ethnicity and Gender, and Standardized Test Performance*, 34 J. APPLIED SOC. PSYCHOL. 665, 665 (2004) (finding no statistical or practical significance).

In this study, the risk of stereotype threat appears to be minimal, given the noncompetitive, nonevaluative nature of the subject matter of the experiment. In addition, it may be instructive to compare the results from Schwartz and Seaman's experiment with those of this study, which uses a slightly modified version of the hypothetical and obviousness questions from Schwartz and Seaman's experiment. Specifically, Schwartz and Seaman's study asked for the respondents' demographic information at the *end* of the survey and their data showed a highly statistically significant effect on obviousness decisions based on gender. Schwartz & Seaman, *supra* note 43, at 452, 479. In the current study, the gender question was presented at the *beginning* of the survey, and the data similarly revealed a highly statistically significant effect based on gender. *See infra* Appendices 3–6. With respect to race, Schwartz and Seaman's study did not report a statistically significant effect, Schwartz & Seaman, *supra* note 43, at 479, while in the current study, race was significant only at the 10% level. *See infra* Appendices 3–6. As mentioned above, this may be partly attributable to the much larger sample size used in the current study (*N* = 1745) compared to that of Schwartz and Seaman's study (*N*=500). *See supra* note 122 and accompanying text.

decisions,[124] which confirms a similar finding by Schwartz and Seaman.[125] With respect to personal experiences relating to patents, only one respondent had served as a juror in a patent case; nine had either applied for or owned a patent; and twenty-five had work experience, expertise, or training in patent law. In total, only thirty-five respondents out of 1745 (2%) had any personal experience relating to patents. Given its relative rarity, patent experience was not included as a variable in the regression models because any indication (or absence) of statistical significance was unlikely to be reliable.

With respect to educational background, there was no statistically significant difference in the odds of an invalidity decision depending on whether the mock juror was a college graduate (1035 out of 1745; 59%) or had a degree (college or graduate) in science, engineering, or mathematics (316 out of 1745; 18%).[126] This confirms a similar finding by Schwartz and Seaman.[127] There were twenty-nine respondents (1.7%) who attended (or were currently attending) law school. A "Law School" variable was not included in the regression models because of reliability concerns arising from the low cell count.

Regarding jury service, 263 respondents (15%) had previously served on a jury, of which ninety-five served in a civil case, 132 served in a criminal case, twenty-one served in both types of cases, and fifteen were unsure of the type of case in which they served. By way of comparison to the overall population, approximately a quarter of adults in the United States have served on a jury.[128] The lower jury service percentage in the dataset might be an artifact of the respondents' median age being only thirty. Indeed, the respondents who had served on juries were older (median/mean age: 40/41.5) than those who had not (median/mean age: 29/32.7) by a highly statistically significant margin.[129] As a further comparison, the median age of individuals

124. *See infra* Appendices 3–6.

125. Schwartz & Seaman, *supra* note 43, at 479–80.

126. *See infra* Appendices 3–6.

127. Schwartz & Seaman, *supra* note 43, at 479–80.

128. Regina A. Corso, *Just Under Three in Five Americans Believe Juries Can Be Fair and Impartial All or Most of the Time*, HARRIS INTERACTIVE, (Jan. 21, 2008), http://www.harrisinteractive.com/vault/Harris-Interactive-Poll-Research-Just-Under-Three-in-Five-Americans-Believe-Juries-2008-01.pdf.

129. According to a t-test assuming unequal variances, the difference in the means is

who were empaneled as jurors in the previously-mentioned 2004 study in King County, Washington, was forty-eight.[130] For these reasons, the results relating to the possible impact of jury service should be evaluated cautiously. With the foregoing caveat in mind, the regression models reveal no statistically significant difference in the odds that a respondent would find invalidity based on the mere fact of prior jury service.[131] However, when comparisons were made based on the type of jury service, the respondents who had served on a civil jury[132] were less likely to find invalidity than someone who either had served on a criminal jury or had not served at all, by a margin that was significant at the 10% level.[133] In contrast, service on a criminal jury[134] did not have any statistically significant effect on the outcome.[135] One possible explanation for these results might be that the 10% significance level for civil jurors is an artifact of the sample size. Another potential explanation might be that former civil jurors who likely have had prior experience applying the preponderance standard might have taken a more rigorous view of the clear and convincing standard (compared to the other respondents) upon being informed in the hypothetical that the latter standard is higher than the former. In contrast, the respondents who have not served on a civil jury may not have a preexisting set point with respect to the preponderance standard such that their conceptual threshold of whether the clear and convincing standard was satisfied might have been lower or more fluid.

In summary, the strongest predictors of whether an individual juror would find invalidity in this study were the treatment scenarios and gender. The gender predictor should be interpreted with caution because of the potentially gendered nature of the subject matter of the hypothetical patent-in-suit (golf balls). The effects of race and prior service on a civil jury were significant

highly statistically significant: $p < 0.0001$.

130. GASTIL ET AL., *supra* note 77, at 61, tbl.4.2.

131. As represented by the "Jury Service" variable in the regression models. *See infra* Appendices 3–6.

132. As represented by the "Civil Juror" variable in the regression models. *See infra* Appendices 3–6.

133. Across the various models, the *p*-value ranged between 0.080 and 0.083. *See infra* Appendices 3–6.

134. As represented by the "Criminal Juror" variable in the regression models. *See infra* Appendices 3–6.

135. *See infra* Appendices 3–6.

only at the 10% level, and no statistically significant relationships could be discerned for age, education, service on a criminal jury, and experience related to the subject matter of the hypothetical patent-in-suit. However, in light of the absence of group deliberation in this study, caution is warranted before drawing any firm conclusions about the relative impact (or lack thereof) of any of these variables on case outcomes.[136]

IV. IMPLICATIONS

The results of this study confirm, in large part, the perceptions of litigants and trial judges on the likely impact of informing the jury of the presumption of validity and presenting information critical of the PTO during trial. Specifically, the data show that the presumption instruction and criticisms of the PTO each have statistically significant effects on the rate at which the mock jurors found the hypothetical patent-in-suit invalid for obviousness.[137] On the whole, each of the treatments appears to have an effect on the invalidity rate that may be tantamount to an incremental, yet material, change in the standard of proof for invalidity.

The impact of the PTO criticisms on the results arguably supports the prevailing view among judges that such information may be highly prejudicial.[138] Because the overwhelming majority of mock jurors did not have any personal experience with the patent system,[139] it is possible that they were highly susceptible to being influenced by *any* information on the PTO, whether positive or negative. The data do not reveal the extent to which the impact of the PTO criticisms might have been attributable to the jurors' susceptibility to being influenced based on their lack of familiarity with the agency, as opposed to the negativity in the message imparted by the criticisms. Because only 2% of the mock jurors had any personal experience with the patent system, it was not possible to determine reliably from the data whether they might have been less affected by the criticisms than those

136. *See infra* notes 158–60 and accompanying text.

137. *See supra* Part III.A.

138. *See supra* note 37 and accompanying text.

139. Of the 1745 respondents, only thirty-five (2%) had any experience with the patent system.

who had no familiarity with the agency. Accordingly, future research may explore whether and to what extent a relationship may exist between the jurors' level of familiarity with the patent system and the degree to which their decisions may be influenced by the presentation of related information that is either positive, negative, neutral, or some combination thereof.

With respect to the effect of the presumption instruction, the optional nature of instructing jurors on the presumption of validity under current Federal Circuit law[140] introduces the possibility for different outcomes on validity issues depending on whether a presumption instruction was given. This would be a concern primarily in close cases, which generally are the ones that go to trial.[141] If some district courts consistently instruct the jury on the presumption of validity, while others do not, the presumption instruction could become another factor in the calculus of forum shopping in patent litigation.[142] In light of these concerns, the district courts might adopt a norm of either including the presumption instruction in every case or excluding it. The difficulty, however, lies in selecting a norm.

From a procedural standpoint, one possible justification for adopting a norm of including the presumption instruction may be that, to the extent it sends a strong signal that patents should not be invalidated lightly, it may help reinforce the message to the jury that the clear and convincing standard associated with deciding validity issues is more rigorous than the preponderance standard associated with deciding other issues in the case, such as infringement. Relatedly, given that patent trials are complex proceedings where the jury is presented with evidence and arguments on multiple topics, the presumption instruction may serve as a procedural safeguard that decreases the likelihood that

140. *See* Chiron Corp. v. Genentech, Inc., 363 F.3d 1247, 1258–59 (Fed. Cir. 2004).

141. *Cf*. George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 17 (1984) ("In litigation, as in gambling, agreement over the outcome leads parties to drop out.").

142. *See generally* Kimberly A. Moore, *Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?*, 79 N.C. L. REV. 889, 892 (2001) ("The empirical results presented in this Article demonstrate that despite the creation of the Federal Circuit, choice of forum continues to play a critical role in the outcome of patent litigation.").

the jury will decide validity issues based on considerations other than the pertinent evidence.[143]

Conversely, other procedural considerations may support the contrary norm of excluding any mention of the presumption. For example, the presumption instruction may be confusing to lay jurors who might mistakenly accord it evidentiary weight or otherwise misinterpret it.[144] In addition, one possible explanation for the disparity in the invalidity rates between the *No Treatments* and the *Presumption Only* scenarios might be that the presumption unfairly discourages jurors from second-guessing the PTO, over and above the level of deference built into the clear and convincing standard for proving invalidity.[145] Indeed, omitting the presumption instruction may allow the standard of proof to be effectively recalibrated—without requiring any change in the law—to a lower level that may better reflect the realities of the examination process at the PTO that have raised concerns about patent quality.[146]

By deciding to include (or exclude) the presumption instruction, a district court is effectively selecting a validity baseline, which may affect the extent to which the clear and convincing standard becomes more (or less) difficult to satisfy. Because the choice of a validity baseline reflects, in part, a normative judgment concerning whether invalidating a patent should be made easier or more difficult, it should not rest solely on the procedural benefits associated with each option. Ideally, the respective error costs should also be considered: A critical characteristic of each baseline is the relative level of false positives and false negatives. In the present context, a false

143. This is somewhat analogous to the role played by an instruction on the presumption of innocence in a criminal trial, where the presumption acts as a safeguard against "a genuine danger that the jury would convict . . . on the basis of . . . extraneous considerations, rather than on the evidence introduced at trial." Taylor v. Kentucky, 436 U.S. 478, 488 (1978).

144. *See supra* note 29.

145. *See, e.g.*, Bohrer, *supra* note 17, at 282–83.

146. *See, e.g.*, Lichtman & Lemley, *supra* note 1, at 49 (suggesting that presumption of validity be changed "to more accurately reflect the realities of current patent practice"); *see also* FTC Report, *supra* note 66, at 8 ("[T]he PTO is underfunded, and PTO patent examiners all too often do not have sufficient time to evaluate patent applications fully. These circumstances suggest that an overly strong presumption of a patent's validity is inappropriate. Rather, courts should require only a 'preponderance of the evidence' to rebut the presumption of validity.").

positive, also called a "Type I" error, occurs if a valid patent is erroneously invalidated.[147] A false negative, which is a "Type II" error, occurs when the court fails to invalidate a patent with a validity defect.[148]

Looking at the two baseline options at a high level, the incidence of Type I errors is expected to increase (or decrease) when the presumption instruction is omitted (or included), in light of the results[149] suggesting that this baseline would effectively impose a lower (or higher) standard of proof than the alternative. For Type II errors, the inverse relationship may hold. Given the infeasibility of eliminating either type of error, the selection of a baseline would necessarily be informed by a need to strike the optimal balance of Type I and Type II errors. Presently, differing opinions exist as to which type of error should be prioritized for avoidance. Some commentators, drawing on comparative error analysis from antitrust law,[150] suggest that Type I errors may be more problematic than Type II errors on the theory that the former is less amenable to correction through market forces than the latter.[151] In addition, a high Type I error rate is thought to introduce a level of uncertainty in patent protection that could dissuade some patentees—particularly those in the pharmaceutical industry—from investing in research and commercialization.[152] By contrast, other commentators have

147. *See* Christopher A. Cotropia, *Patent Law Viewed Through an Evidentiary Lens: The "Suggestion Test" as a Rule of Evidence*, 2006 BYU L. REV. 1517, 1564 n.271.

148. *Id*.

149. *See supra* Part III.A.

150. Error analysis in antitrust law classifies a Type I error as behavior wrongly classified as illegal and a Type II error as a monopoly that is wrongly permitted. *See* Fred S. McChesney, *Talking 'Bout My Antitrust Generation: Competition for and in the Field of Antitrust Law*, 52 EMORY L.J. 1401, 1412–13 (2003).

151. *See, e.g*., Kevin D. McDonald, *Hatch-Waxman Patent Settlements and Antitrust: On "Probabilistic" Patent Rights and False Positives*, 17 ANTITRUST ABA, at 68, 74 (2003) ("Unlike Type II error . . . Type I error in the patent system cannot be corrected."); *see also* McChesney, *supra* note 150, at 1413 ("Type I error . . . is not subject to much self-correction. If liability is imposed on conduct that actually is beneficial (that is, competitive innocents are punished), there is no market corrective for judicial mistake. Only judicial reversal of the case or legislative intervention to change the decision will undo the Type I error."); Frank H. Easterbrook, *On Identifying Exclusionary Conduct*, 61 NOTRE DAME L. REV. 972, 977 (1986) ("[F]alse positives are much more harmful than false negatives. Market processes undercut monopolies wrongfully permitted, but no similar processes undercut judicial decisions that wrongly condemn efficient conduct.").

152. *See, e.g.*, Lichtman & Lemley, *supra* note 1, at 52 (summarizing patentees' rationale for strong presumption of validity).

observed a weak causal relationship between patents and innovation (at least in some industries),[153] and have concluded that Type II errors are more problematic on the basis that they impede follow-on innovation and produce pricing distortions.[154] A detailed analysis that quantifies and compares the relative harms of each error type for the purposing of informing a normative preference for a particular validity baseline is left for future research.

At bottom, in every trial involving a validity issue, a validity baseline specific to that case will be established based on the information presented to the jury. The decision to include or exclude certain information that may materially affect the baseline is a policy-based assessment that should be informed not only by procedural considerations but also by the relative costs of erroneously invalidating patent claims versus erroneously upholding them. Although the results of this study cannot definitively answer the question of which baseline should be chosen among the various treatment configurations, it highlights the need for further empirical research on validity error costs that may allow a court that is faced with a choice between two (or more) legally-permissible procedural options to choose the one that optimizes the tradeoff between Type I and Type II errors.

153. *See, e.g.*, Stuart J.H. Graham et al., *High Technology Entrepreneurs and the Patent System: Results of the 2008 Berkeley Patent Survey*, 24 BERKELEY TECH. L.J. 1255, 1283 (2009) ("Given that the patent monopoly is most commonly justified on the ground of providing incentives to innovate, we were surprised to find that, in general, the technology startup executives responding to our survey report that patents offer relatively mixed to weak incentives to engage in innovation."); FTC REPORT, *supra* note 66, at ch. 3, 33–38.

154. *See, e.g.*, Rochelle Cooper Dreyfuss, *Nonobviousness: A Comment on Three Learned Papers*, 12 LEWIS & CLARK L. REV. 431, 435 (2008) ("Type II errors take material out of the public domain, increase patent thickets and transaction costs, act as barriers to entry and to cumulative research, and encourage trolling. But while type I errors may be bad for the inventor, they can be very advantageous to society."); Ian Ayres, & Paul Klemperer, *Limiting Patentee's Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies*, 97 MICH. L. REV. 985, 1019-20 (1999). ("Enforcing invalid patents creates ex post pricing distortions without enhancing innovation, while our model showed that failing to enforce otherwise-valid patents could reduce the ex post distortions without reducing, or without substantially reducing, innovation incentives.").

V. LIMITATIONS OF THIS STUDY

There are several aspects of this study that may limit its external validity.

First, a notable difference between the survey flow and the order of proceedings in some courts is that, at the beginning of the trial, the jury will be shown an informational video about the PTO that is produced by the Federal Judicial Center ("FJC").[155] The FJC video has received mixed reviews from practitioners, some of whom believe the video helps the patentee more than the accused infringer.[156] The FJC video was not used in this study because it would have more than doubled the length of the survey (the video itself is eighteen minutes long), and overly long surveys may adversely affect the number and quality of respondents. Because the presumption of validity is grounded in the concept of administrative correctness, the content of the FJC video may be relevant to the extent that it provides an extended discussion of the process of obtaining a patent at the PTO. Future research could explore whether the relationships reported in this article would still hold if the FJC video were shown to mock jurors.

Second, the mode of presenting information to the mock jurors is substantially different from real life. During actual trials, information is presented live to the jury in both audible and visual forms, which may affect the retention and salience of certain information.[157] For the survey experiment, the hypothetical was provided as a text document because the survey was conducted online and one of the goals was to minimize any considerations that would discourage participation. Presenting the hypothetical as a video might have been more realistic; however, it might have limited the respondent pool to those individuals who had both access to sufficient bandwidth and the opportunity to watch a video that would have been several

155. *See, e.g.*, John D. Gilleland, *The Debate Is On: Is the Federal Judicial Center's Patent Tutorial Video Too Pro-Plaintiff?*, TrialGraphix 2 (May 1, 2012), *available at* http://www.trialgraphix.com/SiteAssets/file/Articles/patent-tutorial-video-too-proplaintiff-john-gilleland.pdf.

156. *See id.* ("Defense teams claim the video is too pro-plaintiff in that it dedicates a good chunk of its running time to extolling the virtues of the patent system . . . .").

157. Schwartz & Seaman, *supra* note 43, at 470–71.

minutes long. In addition, because individuals who respond to Internet surveys often do so while at work, presenting the hypothetical as a video might have dissuaded some respondents from participating if they wanted to avoid getting caught watching videos at work.

Third, juries deliberate in groups, not individually, such that the decisions of many separate individuals might not be representative of a decision reached by a group.[158] For example, research suggests that the perceived influence of women and minorities in jury deliberation may be lower than that of white men.[159] Moreover, the unanimity requirement for jury verdicts[160] may have a significant influence on group deliberation that may not be reflected in the results of this study. To explore these issues further, future research could use group deliberation for mock jurors with varying demographic compositions.

Fourth, unlike a real trial, which presents a substantial amount of information to the jury over several days,[161] the relative brevity of the experiment[162] may enhance the prominence of the treatments, which may skew the results to show a greater effect than may be possible under real-life circumstances.[163] Indeed, the hypothetical is not representative of the level of complexity in a typical patent trial. Juries in actual patent trials are usually asked to render a verdict on multiple issues (e.g., infringement, invalidity, remedies), often for multiple claims (if not multiple patents), as opposed to only a single issue (obviousness) for a single patent as in this study.[164] As such,

158. *See id*. at 471.

159. *See, e.g.*, Carol J. Mills & Wayne E. Bohannon, *Juror Characteristics: To What Extent Are They Related to Jury Verdicts?*, 64 JUDICATURE 22, 28–29 (1980); *see also* Nancy J. King, *Postconviction Review of Jury Discrimination: Measuring the Effects of Juror Race on Jury Decisions*, 92 MICH. L. REV. 63, 98–99 (1993) (summarizing research relating to influence of minority jurors).

160. Patent cases are heard in federal court, 28 U.S.C. § 1338(a) (2012), in which jury verdicts must be unanimous. FED. R. CIV. P. 48(b) ("Unless the parties stipulate otherwise, the verdict must be unanimous and must be returned by a jury of at least 6 members.").

161. *See* Mark A. Lemley, Jamie Kendall & Clint Martin, *Rush to Judgment? Trial Length and Outcomes in Patent Cases*, 41 AIPLA Q.J. 169, 177 (2013) (reporting that patent jury trials take on average 8.6 days).

162. On average, the mock jurors took 12.3 minutes to complete the survey.

163. *See* Jason Barabas & Jennifer Jerit, *Are Survey Experiments Externally Valid?*, 104 AM. POL. SCI. REV. 226, 227 (2010).

164. *E.g.*, Wendy Kaufman, *Jury to Decide Apple's Patent Case Against Samsung*, NPR (Aug. 22, 2012, 4:50 AM) http://www.npr.org/2012/08/22/159679099/jury-to-decide-apple-s-

concerns about juror confusion are common in patent cases, given the number of issues being presented.[165] For this reason, in a real trial, it is possible that the impact of the presumption instruction and the criticisms about the PTO may be far less salient than the results from the experiment suggest. Future research could explore the impact of the treatments using longer hypotheticals that ask mock jurors to decide multiple issues.

Finally, even if the hypothetical were presented in a more realistic manner, the mock jurors would still be aware that their decisions would not have real-world consequences, which may affect the results.[166] For future research, an analysis of actual jury verdicts where invalidity was decided may help avoid issues related to "the consequentiality of the task" in using simulations.[167] One caveat with analyzing actual jury data is that the complexity and variations among real-life cases may render it difficult to reliably identify and isolate the impact of the presumption instruction on case outcomes.

CONCLUSION

The results of the survey experiment reported in this article suggest that informing the jury of the presumption of validity, as opposed to not mentioning the presumption, could have a substantial impact on the jury's decision on invalidity issues. Specifically, the presence or absence of the presumption instruction may have an effect that is comparable to an incremental change in the standard of proof for invalidity. The results also suggest that similar effects in the opposite direction may occur when jurors are exposed to criticisms of the PTO. Because a jury instruction on the presumption is optional under Federal Circuit law so long as the jury is informed of the clear and convincing standard for proving invalidity, the effect of the

patent-case-against-samsung (reporting that, in light of case complexity, judge "worries that the jury will be 'seriously confused.'").

165. *See id.*

166. *See, e.g.*, David L. Breau & Brian Brook, *"Mock" Mock Juries: A Field Experiment on the Ecological Validity of Jury Simulations*, 31 LAW & PSYCHOL. REV. 77, 89 (2007) (reporting results of experiment finding that "mock jurors might be less emotionally invested in their task than real jurors" and that "this translated into completely opposite verdicts from almost identical trials, apparently stemming from the fact that one jury believed the consequences of its decision were real while the other knew they were not").

167. *Id*. at 80 (internal quotations and citation omitted).

presumption instruction raises concerns about forum shopping, particularly in the absence of a consensus among judges on whether the presumption instruction should be included. A consensus is unlikely to be reached in the absence of further research that quantifies the relative error costs associated with including or omitting the presumption instruction.

APPENDIX 1: HYPOTHETICAL

OVERVIEW OF THE DISPUTE

This dispute is between Acme Golf, Inc., and Bravo Sporting Goods Company. Acme and Bravo are competing manufacturers of golf balls.

Historically, golf balls consisted of two parts: a solid core and a cover-layer with dimples. Some balls had a relatively hard plastic cover-layer because they were designed to travel long distances when struck by a club. However, this hard cover created an undesirable "feel" when struck, and made it difficult for some golfers to control the ball's direction or spin. In contrast, other balls had a soft cover-layer made of polyurethane in order to provide the proper "feel" when struck and greater control for shorter shots. But soft-cover balls had the disadvantage of travelling less distance than their hard-cover counterparts, and were less durable. Both hard-cover and soft-cover balls were well known in the field since at least the 1950s.

In 2005, Acme designed a three-piece golf ball with: (1) a solid core, (2) a hard inner layer, and (3) a softer outer cover-layer of polyurethane covered with dimples. This three-piece design resulted in a "dual personality" ball capable of traveling long distances due to the hard inner layer, but also had the desirable control and "feel" characteristics of soft-cover balls due to the polyurethane cover-layer. Acme timely applied for a patent on this three-piece golf ball in 2005.

In the United States, patents are granted by the U.S. Patent & Trademark Office, also known as the PTO, which is an agency of the federal government. To obtain a patent, one must first file a patent application with the PTO. A technically-trained patent examiner then reviews it to determine whether the claimed invention is patentable. During this process, the patent examiner searches for and reviews certain information called "prior art," which is any publicly-available information about the technology existing before the date the patent application was filed. The patent examiner reviews the "prior art" to determine whether the claimed invention is truly an advance over existing technology.

One requirement for obtaining a patent is that the invention is not "obvious" in light of the prior art. A claimed invention is "obvious" if an ordinary-skilled person in the relevant field of technology—who was familiar with the prior art—would have also been able to come up with the invention at the time the invention was made.

In this case, the patent examiner reviewed the prior art regarding both hard- and soft-cover golf balls. The prior art the patent examiner reviewed included a patent granted to an inventor named Charles in 2000—which is five years prior to when Acme invented its golf ball. The prior art Charles patent discloses a three-piece golf ball with a solid core, a hard inner layer, and an outer cover-layer consisting of a very hard resin covered with dimples. This hard resin surface had the advantage of making the golf ball extremely durable. The Charles patent does not mention polyurethane, nor does it suggest trying to use a softer material for the outer cover-layer of the ball. After reviewing the prior art, including the Charles patent, the patent examiner determined that Acme's three-piece golf ball was not obvious and allowed a patent to be issued to Acme.

Earlier this year, Acme sued Bravo for selling golf balls that allegedly infringe Acme's patent. In response, Bravo has asserted that Acme's patent is invalid for obviousness in light of the prior art—that is, the technology already in existence at the time Acme invented its golf ball. Under the patent law, there is no liability for infringement if the invention claimed in a patent would have been obvious.

THE PARTIES' ARGUMENTS

BRAVO'S ARGUMENTS:

Bravo argues that Acme's patent is invalid for obviousness because it merely combines pre-existing items that were already well-known in the prior art. Specifically, Bravo claims that the prior art Charles patent discloses a three-piece golf ball with inner and outer layers of different hardness. It would have been obvious to an ordinary golf ball manufacturer, Bravo contends, to modify the Charles three-piece ball to have a soft, outer cover-layer of polyurethane, which has been widely used in traditional two-piece soft-cover balls since the 1950s. Because of this

polyurethane cover, a golf ball maker would expect such a ball to have the desirable control and "feel" characteristics of soft-cover balls. Bravo argues that the jury should not defer to the patent examiner's conclusion that the Acme three-piece ball was patentable because the patent examiner did not thoroughly analyze the prior art. In short, Bravo argues that the patent examiner made a mistake in allowing Acme's patent to issue.

[i][[That the patent examiner might have made a mistake should not be surprising. As recognized by numerous academic researchers, poor patent quality is a serious problem. The PTO is underfunded and has a backlog of approximately 600,000 patent applications that are awaiting examination. The patent examiners are overworked, and are simply not given enough time to review patent applications thoroughly. Indeed, according to one academic study, about half of all patents that are litigated in court are found to be invalid.]]

ACME'S ARGUMENTS:

Acme argues that its patent is not obvious for several reasons. Acme asserts that none of the prior art discloses the combination of items that resulted in the patented invention. Acme contends that this combination is worthy of a patent because it creates a golf ball with the unique benefits of the control and "feel" of a two-piece soft-cover ball, combined with the long distance of a hard-cover ball. Acme insists that nothing in the prior art suggests that this combination would create a ball with these favorable characteristics. According to Acme, the prior art Charles patent does not make Acme's patented invention obvious because the golf balls in the Charles patent were designed to solve a very different problem—the lack of durability. Acme further claims that there is nothing in the Charles patent that would suggest to an ordinary golf ball manufacturer that using a softer cover like polyurethane on a three-piece ball might be a good idea. Finally, Acme argues that the jury should defer to the decision of the technically-trained patent examiner, who was in

[i] This is the treatment block that contains criticisms about the PTO. This block was present in the following versions of the hypothetical: *PTO Criticisms Only* and *Both Treatments*.

the best position to determine whether Acme's claimed invention was obvious.

INSTRUCTIONS FOR DECIDING OBVIOUSNESS

There are several rules you must follow in deciding whether Acme's patent is invalid for obviousness. The fact that the PTO grants a patent on a claimed invention does not necessarily mean that it in fact deserves protection under the patent laws. A party can argue in court that it is not liable for infringement because the patent is invalid. Here, Bravo is arguing that Acme's patent is invalid on the ground that the patent examiner made an error in determining that Acme's invention was not obvious.

[ii][[Under the law, Acme's patent is presumed to be valid. In other words, it is presumed to have been properly granted. When a party attacking the validity of a patent relies on prior art that was specifically considered by the patent examiner, that party bears the burden of overcoming the deference due a qualified government agency official who is presumed to have performed his or her job correctly. The presumption of validity that is accorded a duly-issued patent can be overcome by "clear and convincing" evidence of obviousness. In other words,]] [[I]]n[iii] order to prevail, Bravo must persuade you that the claimed invention in the Acme patent is obvious by "clear and convincing" evidence.

"Clear and convincing" evidence means that it is highly probable that a factual assertion is true. This is a higher standard of proof than a "preponderance of the evidence," which means "more probable than not." However, "clear and convincing" evidence is lower than the "beyond a reasonable doubt" standard used in criminal cases.

An invention is "obvious" if a person of ordinary skill in the relevant technical field—who knew about the prior art and the state of technology that existed at the time the invention was

[ii] This is the treatment block that mentions the presumption of validity. This block was present in the following versions of the hypothetical: *Presumption Only* and *Both Treatments*.

[iii] When the presumption treatment block was present, this word was spelled "in." Otherwise, it was spelled "In."

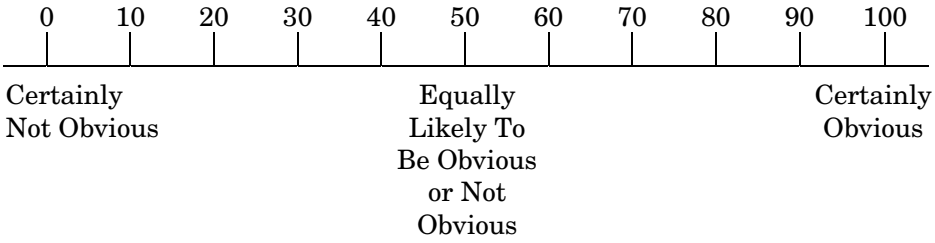
made—would have also come up with the invention at that time. In deciding obviousness, you must avoid using hindsight; that is, you should not consider what is known today or what was learned from the teachings of Acme's patent. In addition, you should not use Acme's patent as a road map for selecting and combining items of prior art.

APPENDIX 2: OBVIOUSNESS QUESTIONS[iv]

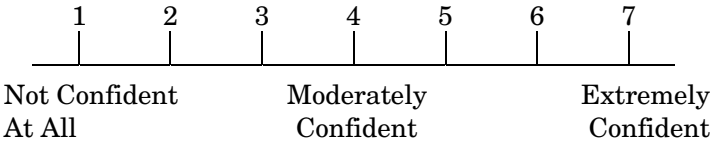
B1. In your opinion, did Bravo prove by clear and convincing evidence that Acme's patent was obvious?

__ Yes (Obvious)
__ No (Not Obvious)

B2. On a scale of 0% to 100%, how likely do you think it is that Acme's patent was obvious?



B3. On a scale of 1 to 7, how confident are you in your answers to the previous two questions (Questions B1 and B2)?



[iv] The questions were adapted from Schwartz and Seaman's study, with slight modifications in the wording and order. *See* Schwartz & Seaman, *supra* note 43, at 478.

APPENDIX 3: LOGISTIC REGRESSION MODELS A1–A4[v]

Variable	Model A1	Model A2	Model A3	Model A4
No Treatments	---	---	---	---
Presumption Only	.7022* (.1071)	.7012* (.1069)	.6965* (.1063)	.6975* (.1065)
PTO Criticisms Only	1.3604* (.1947)	1.3633* (.1951)	1.3717* (.1964)	1.3686* (.1960)
Both Treatments	.9360 (.1372)	.9376 (.1375)	.9346 (.1372)	.9329 (.1370)
Age	.9935 (.0048)	.9933 (.0048)	.9942 (.0048)	.9944 (.0048)
Female	.7157** (.0779)	.7144** (.0780)	.7115** (.0777)	.7125** (.0776)
Minority	1.2582† (.1581)	1.2558† (.1582)	1.2723† (.1605)	1.2753† (.1605)
Golf Player	.9869 (.1076)	.9832 (.1070)	.9865 (.1074)	.9902 (.1080)
College Graduate	.9426 (.1007)	---	---	.9445 (.1011)
Science Degree	---	1.0015 (.1361)	1.0091 (.1373)	---
Jury Service	1.0215 (.1560)	1.0189 (.1558)	---	---
Civil Juror	---	---	.6676† (.1543)	.6701† (.1547)
Criminal Juror	---	---	1.1885 (.2234)	1.1929 (.2241)
Constant	.6693 (.1393)	.6536 (.1347)	.6384 (.1318)	.6542 (.1363)
Prob > Chi2	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Pseudo R^2	0.0173	0.0172	0.0190	0.0191

[v] In Models A1–A4, the *No Treatments* scenario is the base comparison variable to which the other treatment scenarios were compared. Odds ratios are reported with (standard errors). Significance levels: †: $p < 0.10$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

APPENDIX 4: LOGISTIC REGRESSION MODELS B1–B4[vi]

Variable	Model B1	Model B2	Model B3	Model B4
No Treatments	1.4241* (.2171)	1.4262* (.2174)	1.4357* (.2191)	1.4336* (.2188)
Presumption Only	---	---	---	---
PTO Criticisms Only	1.9374*** (.2908)	1.9444*** (.2917)	1.9694*** (.2960)	1.9621*** (.2951)
Both Treatments	1.3329† (.2047)	1.3372† (.2052)	1.3417† (.2062)	1.3375† (.2057)
Age	.9935 (.0048)	.9933 (.0048)	.9942 (.0048)	.9944 (.0048)
Female	.7157** (.0779)	.7144** (.0780)	.7115** (.0777)	.7125** (.0776)
Minority	1.2582† (.1581)	1.2558† (.1582)	1.2723† (.1605)	1.2753† (.1605)
Golf Player	.9869 (.1076)	.9832 (.1070)	.9865 (.1074)	.9902 (.1080)
College Graduate	.9426 (.1007)	---	---	.9445 (.1011)
Science Degree	---	1.0015 (.1361)	1.0091 (.1373)	---
Jury Service	1.0215 (.1560)	1.0189 (.1558)	---	---
Civil Juror	---	---	.6676† (.1543)	.6701† (.1547)
Criminal Juror	---	---	1.1885 (.2234)	1.1929 (.2241)
Constant	.4700 (.0994)	.4583 (.0957)	.4447 (.0932)	.4563 (.0968)
Prob > Chi2	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Pseudo R^2	0.0173	0.0172	0.0190	0.0191

[vi] In Models B1–B4, the *Presumption Only* scenario is the base comparison variable to which the other treatment scenarios were compared. Odds ratios are reported with (standard errors). Significance levels: †: $p < 0.10$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

APPENDIX 5: LOGISTIC REGRESSION MODELS C1–C4[vii]

Variable	Model C1	Model C2	Model C3	Model C4
No Treatments	.7351* (.1052)	.7335* (.1049)	.7290* (.1044)	.7307* (.1046)
Presumption Only	.5162*** (.0775)	.5143*** (.0772)	.5078*** (.0763)	.5097*** (.0766)
PTO Criticisms Only	---	---	---	---
Both Treatments	.6880** (.0991)	.6877** (.0991)	.6813** (.0984)	.6817** (.0984)
Age	.9935 (.0048)	.9933 (.0048)	.9942 (.0048)	.9944 (.0048)
Female	.7157** (.0779)	.7144** (.0780)	.7115** (.0777)	.7125** (.0776)
Minority	1.2582† (.1581)	1.2558† (.1582)	1.2723† (.1605)	1.2753† (.1605)
Golf Player	.9869 (.1076)	.9832 (.1070)	.9865 (.1074)	.9902 (.1080)
College Graduate	.9426 (.1007)	---	---	.9445 (.1011)
Science Degree	---	1.0015 (.1361)	1.0091 (.1373)	---
Jury Service	1.0215 (.1560)	1.0189 (.1558)	---	---
Civil Juror	---	---	.6676† (.1543)	.6701† (.1547)
Criminal Juror	---	---	1.1885 (.2234)	1.1929 (.2241)
Constant	.9105 (.1876)	.8911 (.1820)	.8757 (.1793)	.8954 (.1849)
Prob > Chi2	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Pseudo R^2	0.0173	0.0172	0.0190	0.0191

[vii] In Models C1–C4, the *PTO Criticisms Only* scenario is the base comparison variable to which the other treatment scenarios were compared. Odds ratios are reported with (standard errors). Significance levels: †: $p < 0.10$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

APPENDIX 6: LOGISTIC REGRESSION MODELS D1–D4[viii]

Variable	Model D1	Model D2	Model D3	Model D4
No Treatments	1.0684 (.1567)	1.0665 (.1563)	1.0700 (.1571)	1.0719 (.1574)
Presumption Only	.7502† (.1152)	.7478† (.1148)	.7453† (.1145)	.7477† (.1150)
PTO Criticisms Only	1.4535** (.2095)	1.4541** (.2096)	1.4678** (.2120)	1.4670** (.2119)
Both Treatments	---	---	---	---
Age	.9935 (.0048)	.9933 (.0048)	.9942 (.0048)	.9944 (.0048)
Female	.7157** (.0779)	.7144** (.0780)	.7115** (.0777)	.7125** (.0776)
Minority	1.2582† (.1581)	1.2558† (.1582)	1.2723† (.1605)	1.2753† (.1605)
Golf Player	.9869 (.1076)	.9832 (.1070)	.9865 (.1074)	.9902 (.1080)
College Graduate	.9426 (.1007)	---	---	.9445 (.1011)
Science Degree	---	1.0015 (.1361)	1.0091 (.1373)	---
Jury Service	1.0215 (.1560)	1.0189 (.1558)	---	---
Civil Juror	---	---	.6676† (.1543)	.6701† (.1547)
Criminal Juror	---	---	1.1885 (.2234)	1.1929 (.2241)
Constant	.6264 (.1297)	.6129 (.1260)	.5966 (.1229)	.6103 (.1266)
Prob > Chi2	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Pseudo R^2	0.0173	0.0172	0.0190	0.0191

[viii] In Models D1–D4, the *Both Treatments* scenario is the base comparison variable to which the other treatment scenarios were compared. Odds ratios are reported with (standard errors). Significance levels: †: $p < 0.10$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.