

Diversity, Dialogue, and Deliberation: An Empirical Investigation of Age, Gender, and Meaningful Decision-Making in Korean Juries

*Jisuk Woo and Justin D. Levinson**

I. INTRODUCTION.....	23
II. DIVERSITY AND THE JURY: THEORY AND EMPIRICAL EVIDENCE FROM THE UNITED STATES.....	26
III. BALANCED PARTICIPATION, SUFFICIENT ENGAGEMENT, AND QUALITY OF DELIBERATION.....	29
IV. THE EMPIRICAL STUDY.....	33
A. <i>Data Collection</i>	34
B. <i>Characteristics of the Jurors</i>	36
C. <i>Main Variables and Measurements</i>	37
1. Independent Variables (Diversity and Engagement).....	37
2. Dependent Variables (Quality of Deliberation).....	38
V. RESULTS.....	39
A. <i>Gender and Age Diversity and Quality of Deliberations</i>	39
B. <i>Correlation of the Variables</i>	40
C. <i>Diversity and Quality of Deliberations</i>	41
D. <i>Demographic Diversity and Quality of Deliberations</i>	42
E. <i>Speech Diversity and Quality of Deliberations</i>	43
F. <i>Interaction Effect of Argument Engagement</i>	44
VI. DISCUSSION AND CONCLUSION.....	45

I. INTRODUCTION

American legal discourse has long supported the importance of achieving diversity on juries. Representative juries, in which individuals

* Jisuk Woo is a Professor at the Graduate School of Public Administration, Seoul National University, email: jisuk@snu.ac.kr. Justin D. Levinson is a Professor of Law and the Director of the Culture and Jury Project at the William S. Richardson School of Law, University of Hawai'i at Manoa, email: justinl@hawaii.edu.

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014028257). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the NRF. We thank Professor June Woong Rhee of the Department of Communication at Seoul National University and Jae-Hyup Lee of the School of Law at Seoul National University for consultation on data analysis and helpful comments. Kaitlyn Iwashita provided research assistance.

from a wide range of backgrounds, life experiences, and perspectives participate in deliberation and decision-making, reflect American values of equality, opportunity, and fair treatment for all citizens.¹ The U.S. Supreme Court, reflecting these values, has repeatedly acknowledged that the representativeness of the American jury is as an important element of a litigant's right to a fair trial.²

Scholarly interest in jury diversity in the United States reflects this historical belief that diversity promotes just trial outcomes. Until the past fifty years, however, it was less clear whether and how this belief is realized in jury deliberation processes. Empirical studies have now provided meaningful insight into the connection between representativeness, justice, and participation, and largely support the contention that the best juries possess meaningful diversity.³ Outside of the American context, however, it is unclear whether and how diversity promotes, or perhaps hinders, jury deliberations. This knowledge gap is heightened in non-Western cultural contexts, such as in East Asia. Because "jury-style" systems have emerged in locations such as Taiwan,⁴ Japan,⁵ and Korea,⁶ the cultural context of Western-validated empirical assumptions must be recognized and pursued. Cultural psychological research, for example, has shown differences between Eastern and Western cultures, in areas such as: understanding of causation and intentionality,⁷ conceptions of conflict tolerance,⁸ and views

¹ Eric Y. Cornwell & Valerie Hans, *Representative Through Participation: A Multilevel Analysis of Jury Deliberations*, 45 LAW & SOC'Y REV. 667, 667-698 (2011).

² See *Peters v. Kiff*, 407 U.S. 493 (1972); *Taylor v. Louisiana*, 419 U.S. 522 (1975); *Duren v. Missouri*, 439 U.S. 357 (1979).

³ Cornwell & Hans, *supra* note 1, at 667-68; Nancy Marder, *Juries, Justice, and Multiculturalism*, 75 S. CAL. L. REV. 659, 726 (2002); Samuel Sommers, *On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations*, 90 J. PERSONALITY AND SOC. PSYCHOL. 597, 600-12 (2006).

⁴ See Kuo-Chang Huan & Chang-Ching Lin, *Rescuing Confidence in the Judicial System: Introducing Lay Participation in Taiwan*, 10 J. EMPIRICAL LEGAL STUD. 542 (2013).

⁵ See Kent Anderson & Emma Saint, *Japan's Quasi-Jury (Saiban-in) Law: An Annotated Translation of the Act Concerning Participation of Lay Assessors in Criminal Trials*, 6 ASIAN-PAC. L. & POL'Y J. 233 (2005); Masahito Inouye, *Introduction of the Saiban-in System and Reformation of Criminal Procedure in Japan*, 55 SEOUL L. J. 441 (2014).

⁶ See Jae-Hyup Lee, *Getting Citizens Involved: Civil Participation in Judicial Decision-Making in Korea*, 4 E. ASIA L. REV. 177 (2009); Jae-Hyup Lee, *Korean Jury Trial: Has the New System Brought About Changes?*, 12 ASIAN-PAC. L. & POL'Y J. 58 (2010).

⁷ Justin D. Levinson, *Mentally Misguided: How State of Mind Inquiries Ignore Psychological Reality and Overlook Cultural Differences*, 49 HOW. L. J. 1 (2005).

⁸ Richard E. Nisbett & Takahiko Masuda, *Culture and Point of View*,

regarding the appropriateness of severe punishment.⁹ It has not been studied yet, however, how much these cultural differences may reflect or perhaps challenge the conclusions made by American scholars, including previous findings that: (1) diversity of background adds to the quality of deliberations, and (2) better quality deliberations involve meaningful participation by the most jurors possible.

In this Article, we seek to determine the role of diversity in Korean jury deliberations and contribute to a nascent field of empirical literature focusing on East Asian juries. Furthermore, we investigate whether speech diversity,¹⁰ in addition to age and gender diversity,¹¹ affects the quality of jury deliberations. To pursue these goals, we coded and analyzed data from a unique study¹² of full video recordings of thirty-two shadow juries that sat in for nineteen actual criminal cases in Korea.¹³ The remainder of the Article is structured as follows: Section II begins by considering the diversity of American juries, from a theoretical and empirical perspective. It notes the differing American theoretical perspectives regarding fair cross section and cultural experience, and describes studies documenting the important role of diversity in decision-making. It concludes by contextualizing this American work in Korea, where juries are a new part

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES 11163 (2003).

⁹ Levinson, *supra* note 7.

¹⁰ “Speech diversity” means how evenly each juror spoke during the deliberation.

¹¹ Note that there is little racial and ethnic diversity in Korea. Ethnicity and race are not even factors that are surveyed in the Korean national census. In 2016, 3.9% of the population of Korea were foreign nationals, half of whom were Chinese. *Gugnae Chelyu Oegug-in 200 Manmyeong Dolpa...2021 Nyeon 300 Man Jeonmang* *국내 체류 외국인 200 만명 돌파...2021 년 300 만 전* [Over Two Million Foreigners in Korea... Predicted to Reach Three Million By 2021], YONHAP NEWS (July 27, 2016), <http://www.yonhapnews.co.kr/bulletin/2016/07/27/0200000000AKR20160727075600004.HTML>.

¹² Few empirical studies of juries have had access to real criminal trials. Due to the collaboration between the Supreme Court of Korea and scholars at Seoul National University, a “shadow jury” program was established whereby mock jurors were present for actual trials.

¹³ South Korea instituted a jury system for the first time in its history in January of 2008, when the Act for Civil Participation in Criminal Trials of 2007 was first implemented. The policy goal of this legislation was to increase democratic participation and reinforce the democratic legitimacy of the judicial process, and to enhance the transparency and credibility of the judiciary. For more details about the Korean jury system, see Sangjoon Kim et al. *Judge-Jury Agreement in Criminal Cases: The First Three Years of the Korean Jury System*, 10 J. EMPIRICAL LEGAL STUD. (2013); Jae-Hyup Lee et al. *Analyzing Shadow Jury Deliberations in Korean*, J. KOREAN L. 41 (2013). As in the United States jury system, deliberations are secret, but the Supreme Court of Korea and scholars at Seoul National University collaborated in creating a ‘shadow jury program’ to facilitate empirical research. This article relies upon data collected during shadow jury research.

of the legal system, having been introduced to criminal trials in 2008. Section III turns to examine the role of juror engagement and speech in deliberations. It reviews empirical research from American juries, and describes the different theories of jury deliberations, highlighting what makes these juries most effective. Section IV presents the methods and materials of the empirical study we designed. We conducted an in-depth analysis of demographic and speech diversity in Korean jury deliberations, using full-length videos of thirty-two jury deliberations in nineteen real cases, as part of a program sponsored by the National Court Administration of Korea. Section V presents the results. We found that age and gender diversity harmed the efficacy of Korean jury deliberation, while speech diversity predicted success. Section VI contextualizes these results, looks towards future research, and concludes.

II. DIVERSITY AND THE JURY: THEORY AND EMPIRICAL EVIDENCE FROM THE UNITED STATES

The jury has long been a core component of the American legal system, but scientific interest in jury decision-making processes has emerged only recently. Until the 1950s, the idea of using empirical studies to understand legal assumptions had not yet emerged. Since the 1990s, there has been a relative explosion of empirical methods in legal analysis. Yet, relevant data on jury diversity is still less robust than commentators might hope, partly due to the difficulty researchers face in accessing and analyzing realistic jury deliberations.¹⁴ Nonetheless, because of the long-held American view that the representativeness of diverse groups in the jury is an essential element of the defendant's right to a fair trial, some research has focused on evaluating whether diversity indeed leads to fairer trial outcomes.

Scholars have investigated the effect of jury diversity using a variety of methods. Leslie Ellis and Shari Diamond found, for example, that jury diversity indeed promotes the sense of a fair trial.¹⁵ Ellis and Diamond found that when citizens were provided with the description of a trial involving an all-White jury and a Black defendant, those citizens perceived a guilty verdict to be less fair than when they were told that the verdict was given by a mixed-race jury.¹⁶ Diversity has also been found to promote juror satisfaction with the deliberation and verdict. Nancy Marder, in her research on demographic diversity (gender, race, and age) in deliberations, found that

¹⁴ Dennis J. Devine et al., *Deliberation Quality: A Preliminary Examination in Criminal Juries*, 4 J. EMPIRICAL LEGAL STUD. 273, 304 (2007).

¹⁵ Leslie Ellis & Shari S. Diamond, *Race, Diversity, and Jury Composition: Battering and Bolstering Legitimacy*, 78 CHI.-KENT. L. REV. 1033, 1060 (2003).

¹⁶ *Id.* at 1043-48.

jurors who participated in more diverse juries reported being more satisfied with the deliberations and the verdict.¹⁷

Other empirical studies of jury diversity in the United States have examined how gender plays a role in the deliberation process.¹⁸ An empirical study by Marder on cultural diversity and deliberations found that some types of diversity, such as gender and age, can make a positive difference in the tone and thoroughness of deliberations.¹⁹ Gender had the strongest effect on deliberations.²⁰ As the jury became more gender diverse, the tone became less hostile, more harmonious, and jurors became more supportive of each other.²¹ Furthermore, as gender diversity increased, the jurors perceived the deliberations to be more thorough.²² In addition, jurors were also more satisfied with their deliberations, their jury experience, and their verdict.²³ Marder suggested that courts should strive for petit juries that contain an equal or almost equal number of men and women.²⁴

Studies using mock trials similarly concluded that more diversity resulted in more thorough deliberation. Samuel Sommers found that racially heterogeneous groups (with White and Black jurors) deliberated longer and considered a wider range of information than did racially homogeneous groups (with White jurors only).²⁵ Sommers also pointed out that these differences did not solely result from Black participants adding different perspectives to the discussion, but also because White participants raised more case facts, made fewer factual errors, and were more amenable to the discussion of race-related issues when they were in a diverse group versus in an all-white group.²⁶ Sommers concluded that racially diverse groups were more thorough and competent than homogenous ones: "Jury representativeness can be more than a moral or Constitutional ideal; it is sometimes an ingredient for superior performance."²⁷ This finding suggests that jury diversity is not only important for equal access and legitimacy, but

¹⁷ Marder, *supra* note 3, at 659.

¹⁸ Andrea Hickerson & John Gastil, *Assessing the Difference Critique of Deliberation: Gender, Emotion, and the Jury Experience*, 18 COMMUNICATION THEORY 281, 290-303 (2008); Sommers, *supra* note 3 at 597-612.

¹⁹ Marder, *supra* note 3.

²⁰ *Id.* at 659.

²¹ *Id.* at 701.

²² *Id.*

²³ Marder did not find that race diversity had any effect. *Id.*

²⁴ *Id.* at 703.

²⁵ Sommers, *supra* note 3, at 597-612.

²⁶ *Id.* at 606.

²⁷ *Id.* at 608.

also is an important element that changes the dynamics of the deliberation process. This leads us to think that it would be worth examining the ways in which the diversity of the jury influences the deliberation process more systematically.

Another core element of American research on jury diversity is the size of the jury, in that larger juries potentially allow for more diversity to manifest within a diverse society. Shari Diamond and colleagues analyzed jury selection and composition in 277 civil jury trials and found that jury size had a substantial effect on minority representation.²⁸ In their study, twelve-member juries included more minority representation than six-member juries.²⁹ Diamond concluded that reducing jury size would negatively influence representation of minorities on the jury and criticized contemporaneous reform efforts to reduce the jury size.³⁰ William Schreckhise and Charles Sheldon found in their study on juries in Washington, D.C., that supplementing juror source lists with driver's license and state identification card lists, rather than exclusively using voter registration lists, resulted in greater diversity.³¹

These studies have made important contributions to the literature on jury diversity by analyzing mock trial deliberations using objective criteria. In Korean juries,³² the study we present similarly employs a diversity-based approach, applying concepts of demographic and speech diversity to real trials.

²⁸ Shari S. Diamond et al., *Achieving Diversity on the Jury: Jury Size and the Preemptory Challenge*, 6 J. EMPIRICAL LEGAL STUD. 441 (2009).

²⁹ *Id.* at 435.

³⁰ *Id.* at 427.

³¹ William D. Schreckhise & Charles H. Sheldon, *The Search for Greater Juror Diversity: The Case of the U.S. District Court for the Eastern District of Washington*, 20 JUSTICE SYSTEMS J. 95, 112 (1998).

³² The Korean jury system, implemented in 2008, incorporates elements of both the U.S.-style system and the German lay assessor system. The adoption of the jury system in Korea was mainly driven by participatory democratic concerns. The two main purposes of the Korean jury system are to reinforce the democratic legitimacy of the judicial process and to enhance the transparency and credibility of the judiciary. The Korean jury trial is limited to certain type of criminal cases and is only invoked upon the defendant's choice. Jury verdicts are advisory in nature. The number of jurors varies according to the severity of the case; it can be five, seven, or nine. The more severe the case is, the more jurors are appointed. Jurors discuss the guilt of the defendant in secret and try to reach a unanimous verdict. If the jurors cannot reach a unanimous verdict, they must hear the judges' opinion. After the judges and the jurors have discussed the guilt of the defendant together, the jurors, again without the presence of the judges, enter a verdict based on a simple majority. If the defendant is found guilty, then Korean jurors discuss the sentence with the judge and each juror submits his/her sentencing opinion individually. *See Lee (2010), supra* note 6.

III. BALANCED PARTICIPATION, SUFFICIENT ENGAGEMENT, AND QUALITY OF DELIBERATION

To contextualize our empirical study of diversity and deliberation in Korean jurors, we next consider the dynamics of the deliberation process. We ask: what do the best juries do when deliberating and how can one measure these qualities? In pursuing this inquiry, we seek to further engage empirical studies of jury deliberation by evaluating how particular elements of the deliberations, such as speech diversity and argument engagement, are related to overall jury success.

Empirical studies measuring the quality of deliberations have been plentiful, but still less common than one might expect, despite historically prominent discussions in the literature indicating that dynamics of deliberations can influence their outcomes.³³ In American juries, Reid Hastie and colleagues have led the way in developing a model of jury decision-making, which characterizes two differing, yet dominant, styles of deliberations: verdict-driven and evidence-driven.³⁴ Verdict-driven deliberations occur when jurors announce their position about the case at the beginning of the discussion, followed by citations to evidence in support of that specific position.³⁵ On the other hand, evidence-driven deliberations occur when jurors' opinions are expressed late in the group discussion and evidence is first reviewed without reference to the verdict categories.³⁶ Researchers have found that evidence-driven deliberations generally have higher levels of participation, and include wider-ranging discussions.³⁷ In contrast, verdict-driven deliberations feature early and frequent polling, and harbor pressures to conform to the majority. Interestingly, they also include a greater likelihood of a hung jury.³⁸

Despite the research favoring evidence-driven deliberations, scholars who study American juries found that verdict-driven juries are

³³ REID HASTIE ET AL., *INSIDE THE JURY* (1983); James H. Davis et al., *Order Effects in Multiple Decisions by Groups: A Demonstration With Mock Juries and Trial Procedures*, 47 J. PERSONALITY AND SOC. PSYCHOL. 1003, 1003-12 (1984); John H. Davis et al., *Effects of Straw Polls on Group Decision Making: Sequential Voting Pattern, Timing, and Local Majorities*, 55 J. PERSONALITY AND SOC. PSYCHOL. 918, 918-26 (1988); Shari S. Diamond et al., *Juror Decisions During Civil Trials: Studying an Arizona Innovation*, 45 ARIZ. L. REV. 82 (2003); James A. Holstein, *Jurors' Interpretations and Jury Decision Making*, 9 L. & HUM. BEHAVIOR 83, 90-100 (1985); Sarah Tanford & Steven D. Penrod, *Jury Deliberations: Discussion Content and Influence Processes in Jury Decision Making*, 16 J. APPLIED SOC. PSYCHOL. 322, 322-47 (1986).

³⁴ HASTIE ET AL., *supra* note 33, at 163.

³⁵ *Id.*

³⁶ *Id.*

³⁷ Devine et al., *supra* note 14, at 622-25.

³⁸ HASTIE ET AL., *supra* note 33, at 163.

more common than evidence-driven juries. Furthermore, researchers who have developed their own theories of deliberation have identified a general leniency bias whereby jurors' first ballots tend to predict final verdicts.³⁹ The famous jury researchers Harry Kalven and Hans Ziesel even concluded, on the basis of survey data from 225 actual jury trials, that "the real decision is often made before the deliberation begins."⁴⁰ In this context, jurors who actively participate by making comments and arguments based on evidence make deliberations a meaningful fact-finding process. Despite research showing the importance of jurors' first impressions, other studies have found more optimistic results. For example, Neil Vidmar and Shari Diamond concluded that American jurors indeed pay close attention to the content of the evidence, including expert testimony.⁴¹

In their reflective look at decades of work on jury decision-making, Dennis Devine and colleagues suggested five process-oriented criteria for evaluating the performance of a jury as a whole: instruction comprehension, evidence review, factual focus, systematic participation, and informational influence.⁴² Employing these criteria, Devine and colleagues conducted post-trial surveys from jurors and legal professionals associated with 197 criminal jury trials in Indiana and evaluated their criteria, ultimately reaching mixed conclusions.⁴³ On the positive side, study results showed that a strong majority of the juries reported that they indeed reviewed the evidence thoroughly and understood their instructions. However, less positively, data on systematic participation, maintaining a factual focus, and openness to opinion changes based on informational influence, suggested uneven participation and a verdict-driven deliberation style.⁴⁴

In one of the few empirical jury studies in Korea, Jae-Hyup Lee and colleagues employed Devine and colleagues' criteria to examine the

³⁹ *Id.*; HARRY KALVEN & HANS ZEISEL, *THE AMERICAN JURY* (1966); Robert MacCoun, & Norbert Kerr, *Asymmetric Influence in Mock Jury Deliberation: Jurors' Bias for Leniency*, 54 J. PERSONALITY AND SOC. PSYCHOL. 21, 21-33 (1988); Marla Sandys & C. Dillehay, *First-Ballot Votes, Predeliberation Dispositions, and Final Verdicts in Jury Trials*, 19 L. & HUM. BEHAVIOR 175, 180-95 (1995).

⁴⁰ KALVEN & ZEISEL., *supra* note 39, at 488.

⁴¹ Neil Vidmar & Shari S. Diamond, *Juries and Expert Evidence*, 66 BROOK L. REV. 1121, 1180 (2001). These positive attentional effects may be less prominent in legal cultures like Korea, where jurors are less familiar with the ways that jury trials progress, and thus may not know to focus on expert testimony, for example. When formulating our own study on Korean jurors, we were therefore particularly interested in the ways jurors would talk (or not talk) about evidence during deliberations.

⁴² Devine et al., *supra* note 14, at 273-80.

⁴³ *Id.*

⁴⁴ *Id.*

rationality of jury deliberations of Korean juries.⁴⁵ Lee and his colleagues found that Korean shadow jurors mostly understood the judge's instructions, and understood and recalled the evidence presented at trial. The jurors also paid close attention during the court proceedings and actively participated in deliberations. On the other hand, the degree of evidence-based review seems to have been relatively low, as the jurors tended to vote very early rather than wait until they had fully discussed the facts and evidence of the case. The observations in Lee and colleagues' study suggest that although there is room for improvement in Korean jury deliberations, the overall quality of deliberations was high, considering Korea only introduced the jury system five years before the study. More encouragingly, the deliberations of older or less educated jurors were equal in quality to those of jurors who were younger, more educated, and had a preexisting interest in the jury system. Notably, both the Devine and Lee studies relied on self-reports from jurors, which provided more insight into how the jurors perceived the jury experience, rather than how they actually deliberated. Recognizing the need to build upon self-report studies by employing observational methods, the study we devised focuses on direct observation of each juror's communication behavior to determine the quality of deliberations.⁴⁶

Scholars have also examined how American individual jurors' behavior may influence the overall quality of deliberations. One of the biggest challenges to quality deliberations seems to be passive jurors, who are often completely disengaged from the deliberation process. Diamond and Casper studied when jurors are more likely follow the model of "active information processor" rather than the "jury as a passive participant" by looking at when and how often jurors follow judicial instructions.⁴⁷ They found, as expected, that active jurors are those who participate in evidence interpretation, as well as actively communicate with other jurors.⁴⁸ Sarah Tanford and Steven Penrod build on these concepts by suggesting that further study should include a coding-based examination of dynamics of communication and static content categories.⁴⁹ Dynamics of

⁴⁵ Jae-Hyup Lee et al., *Baesimwon pyeonguiui hapriseonge gwanhan yeongu: gukminchamyeojaepaneseoui geurimja baesim pyeongui bunseok [A Study on the Rationality of the Jury Deliberation: Analysis of the Shadow Deliberations in Korean Jury Trials]*, 139 THE JUSTICE 208 (adding one more criterion to Devine et al.'s five criteria, a cultural aspect of Korean society: exclusion of emotion).

⁴⁶ Devine et al., *supra* note 14, at 273-80; Lee et al., *supra* note 45, at 140.

⁴⁷ Shari S. Diamond & Jonathan D. Casper, *Blindfolding the Jury to Verdict Consequences: Damages, Experts, and the Civil Jury*, 26 L. & SOC'Y REV. 513, 520-557 (1992).

⁴⁸ *Id.* at 513-63.

⁴⁹ Tanford & Penrod, *supra* note 33, at 322-47.

communication, for example, can be observed by measuring several different criteria, such as by examining communications based upon what communications theorists call “two-way” or “multi-way” interactions.⁵⁰ Building upon this scholarship, the study we designed for Korean jurors focuses on the interactivity of juror communication. Specifically, we examined how frequently jurors make comments in response to other jurors’ comments or arguments.

Jürgen Habermas’ idea of a deliberative public sphere is also relevant to assessing what constitutes meaningful deliberation.⁵¹ Deliberative theorists agree that deliberations should stress rational, critical arguments (or, at least, reasoned arguments).⁵² Deliberative theorists and practitioners of the jury system are generally wary of the role of emotion, in contrast to reason, in jury deliberations.⁵³ Their focus on reasoned arguments, however, does not mean that jury deliberations should exclude emotion altogether. Studies have found that emotion deepens deliberations, encourages more participation, and motivates further deliberations.⁵⁴ Andrea Hickerson and John Gastil found, in their survey study of over 3,000 jurors, that experiencing emotion was fairly constant for both men and women. In addition, balancing reason and emotion may be the key for a positive juror experience.⁵⁵

When South Korea was in the process of introducing the jury system for the first time in its history, one of the strongest arguments used to oppose the jury system asserted that Korean people tend to be overly emotional, to the extent that making reasoned arguments and reaching consensus in a rational manner is impossible.⁵⁶ Lee and colleagues’ study of shadow jurors examined how often jurors made overly emotional comments that inhibited

⁵⁰ JOHN FISKE, INTRODUCTION TO COMMUNICATION STUDIES (1990); WILLIAM W. WILMOT, DYADIC COMMUNICATION (1987).

⁵¹ JÜRGEN HABERMAS, THE THEORY OF COMMUNICATIVE ACTION: REASON AND THE RATIONALIZATION OF SOCIETY (Thomas MacCarthy trans., 1985); JUREGN HABERMAS, THE STRUCTURAL TRANSFORMATION OF THE PUBLIC SPHERE (Thomas Burger trans., 1991).

⁵² Joshua Cohen, *Deliberation and Democratic Legitimacy*, in THE GOOD POLITY: NORMATIVE ANALYSIS OF THE STATE 17-34 (Hamlin & Pettit eds., 1989).

⁵³ Simon Thompson & Paul Hoggett, *The Emotional Dynamics of Deliberative Democracy*, 29 POL’Y & POL. 352, 352-364 (2001); Lee et al., *supra* note 45, at 208-30.

⁵⁴ David Michael Ryfe, *The Practice of Deliberative Democracy: A Study of 16 Deliberative Organizations*, 19 POL. COMM. 359, 360-377 (2002); Bas van Stokkom, *Deliberative Group Dynamics: Power, Status, and Affect in Interactive Policy Making*, 33 POL’Y AND POL. 387, 400-09 (1992).

⁵⁵ Hickerson & Gastil, *supra* note 18 at 597-612.

⁵⁶ Sangjoon Kim et al., *Judge-Jury Agreement in Criminal Cases: The First Three Years of the Korean Jury System*, 10 J. EMPIRICAL LEGAL STUD. 35, 35-53 (2013); Lee et al., *supra* note 13, at 43.

the rational discussion necessary for fact-finding. The researchers found that such comments were actually not very common.⁵⁷ A juror's ability to make reasoned arguments and engage in logical thinking is an important element that increases the quality of deliberations. Our study examines how jurors use logic and reason instead of simply arguing their thoughts and positions without any basis besides emotion.⁵⁸

Based on the foregoing theories and empirical findings, we believe that one meaningful way that the quality of deliberations can be assessed empirically is by measuring how much each juror comments on the evidence of the case, responds to other jurors, and uses logic and reason in his or her arguments. We therefore devised a study that employed such measures and attempted to address the following questions in the context of Korean juries:

1. How does age and gender diversity influence the quality of jury deliberations?
2. How does speech diversity influence the quality of jury deliberations?
3. How does the level of argument engagement influence the quality of jury deliberations?

IV. THE EMPIRICAL STUDY

Building upon the theories and empirical findings regarding diversity and quality of jury deliberations in America and Korea, we designed a study to consider the deliberation dynamics that influence the quality of deliberations of Korean juries. First, we examined whether and how age and gender diversity are associated with the quality of deliberations. Next, we investigated whether and how speech diversity, an important parameter of the dynamics of jury deliberations, influenced the quality of Korean jury deliberations. We predicted that, first, diverse juries would demonstrate better performance, and second, that more speech diversity during the deliberation process would be associated with higher quality of jury deliberations. We also explored whether and how argument engagement was associated with the relationship between speech diversity

⁵⁷ Lee et al., *supra* note 45, at 232 (in the study, a juror said, “[t]he poor defendant could not help acting that way due to his childhood experience, so his dad, not the defendant, should be the one to be punished.” Another juror said, “If the defendant goes to jail, what is going to happen to other family members? I feel sorry for him and the family. I do not want to see him punished.” Yet another juror said, “I feel sick that she drank and smoke while she was pregnant. I know that kind of a person. She could have burned her skin with her own cigarette.”).

⁵⁸ We did not include the absence of emotion as an indicator of quality deliberations, as previous research suggested that emotion itself does not necessarily conflict with reasoned arguments. *See generally* Ryfe, *supra* note 54; van Stokkom, *supra* note 54; Hickerson & Gastil, *supra* note 18.

and deliberation quality. In this section, we describe the methods, materials, and results of the study.

A. *Data Collection*

Our methods and statistical analysis draw upon the data we collected for the shadow jury deliberation studies that were conducted with the official cooperation of the National Court Administration of Korea (“NCA”). The NCA coordinated with the courts to set up shadow juries consisting of real juror candidates who were not ultimately selected during voir dire. In addition, a number of district courts in Korea started their own shadow jury programs beginning in September 2010, in order for people to actually experience the judicial process through voluntary participation. These courts intended to improve the public perception of the judiciary’s credibility through this program.⁵⁹ In each trial included in our study, one or two shadow jury groups were formed in addition to the actual jury. Some groups were composed of juror candidates who were summoned by the court to serve as jurors, but were not selected during voir dire (we refer to each such panel as a “venire jury”). Other groups of shadow jurors were composed of people who were directly recruited by the court to serve as shadow jurors (each such panel, a “recruited jury”).⁶⁰

The shadow jurors observed the trial, participated in jury deliberations, and reached a verdict in parallel to the real jury. Shadow jurors were also given the same bench memo (facts, legal issues, etc.) provided to the real jury. After the conclusion of the court proceeding, they were assigned different rooms in the same court, and were engaged in the same process of jury deliberations in secret. When the judge’s intervention was needed, an associate judge of the three-judge panel went into each room to discuss the case with the shadow jurors.⁶¹ The shadow jurors remained until the final court judgment was rendered, frequently late into the night. The research team observed the trials from beginning to end, and then videotaped the shadow jury deliberations. The team did not intervene in the

⁵⁹ The shadow jury program has been successfully implemented. About 16,000 people have participated in the shadow jury program for the past 10 years since the program began on Jan. 1, 2008. Press Release, Supreme Court of Korea, 2017 Nyeon Beob-Won Jeonsigwan Gihoegjeonsi 'Gugminchamyehojaepan' Gaemag 2017 년 법원전시관 기획전시 '국민참여재판' 개막 [Opening of the Shadow Jury Program Planned for the 2017 Court Exhibition] (Sept. 14, 2017), <http://www.scourt.go.kr/portal/news/NewsViewAction.work?pageIndex=1&searchWord=%B1%B9%B9%CE%C2%FC%BF%A9&searchOption=&seqnum=1347&gubun=6>.

⁶⁰ Both venire jurors and recruited jurors were paid KRW 50,000.

⁶¹ All jury trials are conducted by a three-judge panel in Korea. As with the judge’s intervention in the shadow juries, in real trials Korean jurors summon judges in similar situations.

deliberations in any way other than observing and recording the process, in order to preserve the objectivity of the research.

A total of thirty-two shadow juries of nineteen trial cases were set up and operated between January 2012 and June 2012 in cities across the country including Seoul, Ulsan, Gwangju, Jeonju, and Chuncheon. We included both types of shadow jury groups (venire juries and recruited juries) in the study. Through this process, 108 jurors in venire juries and 134 jurors in recruited juries (a total of 242 jurors) participated in the deliberation process (Table 1). The number of jurors in each shadow jury varied from four to twelve members.⁶² Our research team created verbatim transcripts of all the deliberations and conducted a content analysis of the jury deliberations.⁶³ Three coders watched the videos and read the transcripts together until they reached a mutually agreed upon coding, or categorization of data, and then each coder analyzed the videotapes and the transcripts separately.

Table 1: Composition of Juries for Research⁶⁴

		Venire Jury	Recruited Jury	Total
	Area	Number of jurors (Number of jurors in videotape, if different)	Number of jurors (Number of jurors in videotape, if different)	
1	Seoul Central	7	7	14
2	Seoul Central	12	A 11 B 7	30
3	Changwon	9	8	17
4	Cheongju	5	6	11
5	Seoul East	6	8	14
6	Seoul North	8	8	16
7	Ulsan	7	9	16
8	Gwangju	8	7	15
9	Cheongju	4	5	9

⁶² Actual Korean juries consist of five, seven, or nine members. The size of the shadow jury panels deliberately varied from the size of actual juries, as we wanted to examine the effect of different panel sizes on jury deliberations.

⁶³ The length of deliberation ranged from 24 minutes to 104 minutes.

⁶⁴ Not enough potential jurors were left to make a venire jury in one of the trials. Recruited jurors were not recruited in one of the trials. Videotaping was not possible due to the circumstances of the court in four of the trials.

10	Seoul South	0, jury not made	A 7 B 9	16
11	Gwangju	7	4	11
12	Seoul South	8 (0, no videotape)	7	7
13	Seoul North	5	4	9
14	Gwangju	6	9	15
15	Jeonju	5	11(0, no videotape)	5
16	Chuncheon	7	11	18
17	Seoul North	6	0, jury not recruited	6
18	Seoul Central	5 (0, no videotape)	7	7
19	Seoul Central	6	15 (0, no videotape)	6
	Total	121 (108)	160 (134)	242

B. Characteristics of the Jurors

In order to collect the socio-demographic characteristics of the jurors, a survey was conducted after deliberations. Of the 242 jurors, 43.8% were men and 56.2% were women. By age, jurors included persons in their twenties or younger (56.4%), thirties (15.3%), fifties (14.4%), forties (7.2%), and sixties or older (6.8%). One of the reasons for the large number of individuals in their twenties is that many of the recruited jurors were college students. The level of education of the jurors tended to be high, as 52.2% had a college education or higher, 41.8% had a high school education, and only 6.0% received a middle school education or lower.

Table 2: Demographic Characteristics

		Number	Percentage
Gender (not available- 8)	Male	103	43.8
	Female	132	56.2
Age (not available- 6)	20s or younger	133	56.4
	30s	36	15.3
	40s	17	7.2
	50s	34	14.4
	60s and older	16	6.8
Education (not available- 10)	Middle school graduate or lower	14	6.0
	High school graduate	97	41.8
	College graduate or higher	121	52.2

C. Main Variables and Measurements

We used variables at both the jury panel level and the individual juror level, as described below.

1. Independent Variables (Diversity and Engagement)

In order to measure the ‘demographic diversity’ of jury composition, we used the following formula.

$$\text{Demographic Diversity} = - \sum_{i=1}^x (r_i \times \ln(r_i))$$

$r_i = \text{percentage per group}$

We divided jurors into four groups according to gender and age: Males in their twenties and thirties, Males in their forties and older, Females in their twenties and thirties, and Females in their forties and older.⁶⁵ We put natural logarithms on r_i , where r_i was the percentage per group and obtained entropy by calculating \sum . When the calculated value is close to 1, all four groups exist in each jury, indicating higher demographic diversity. When the number is far from 1, the jury consists of only one or two of the groups, indicating a greater similarity in the jurors’ gender and age in the group and thus, less demographic diversity.

In order to measure the ‘speech diversity’ of each jury, which means how evenly each juror spoke during the deliberation, we used the following formula.

$$\text{Speech Diversity} = - \frac{\sum_{i=1}^n (r_i \times \ln(r_i))}{\ln(n)}$$

$r_i = \text{percentage of individual utterances}$
 $n = \text{number of jurors}$

We obtained a value representing entropy by calculating the natural logarithms and \sum of r_i , where r_i was the percentage of individual utterances. A greater entropy value indicates greater uncertainty regarding who will make a comment. A low entropy value is observed when one or two jurors dominate the conversation and it becomes predictable which individual will speak. Conversely, the entropy value is high when many jurors speak, as the uncertainty of who will speak is also high. If the total number of jurors of \sum becomes large, entropy will also be a high value. In order to control the size of the juries in comparing speech diversity, we divided the \sum by the

⁶⁵ Due to the study population consisting of shadow jurors who tended to be young, we did not break down age categorizations further.

number of the jury. Therefore, the speech diversity among the jurors becomes high when the entropy becomes high. The closer this calculated value is to 1, the more evenly speaking is distributed amongst the jurors during the deliberations.

We also measured the level of “argument engagement” by examining how many times a juror’s remarks and other jurors’ replies were exchanged on each issue. The standard by which engagement is measured is whether comments and replies were exchanged more than three times. The level of argument engagement is measured with three levels: 1) no engagement (no reply), 2) medium engagement (three or fewer exchanges), and 3) strong engagement (more than three exchanges).

The values of each of the jury level variables are presented in Table 3.

Table 3: Values of the Jury Level Variables

	N	Minimum	Maximum	Average	SD
Demographic diversity	32	0.00	1.35	0.89	0.29
Speech diversity	32	0.79	1.18	0.92	0.06
Argument engagement	32	1.00	3.00	2.14	0.65

2. Dependent Variables (Quality of Deliberation)

The research team analyzed each juror’s behavior during deliberation. The unit of analysis is every utterance of each juror. An ‘utterance’ is measured as an instance when a juror verbalized a substantive and intelligible sentence lasting at least 5 seconds. A short interjection such as “yes” or “agree” was not coded as an utterance. An unclear expression of opposition such as “uh-” was not coded as an utterance either. On average, each juror made 10.63 utterances, and the maximum number of a single juror’s utterances was 41. The highest number of average utterances of the jurors in a single deliberation was 25.83, and the lowest number of average utterances in a deliberation was 1.55.

We defined the quality of deliberations to depend on the number of meaningful utterances made during the deliberations and the characteristics of such utterances. The quality of deliberations was measured by how often each juror made remarks using supporting evidence, how often the jurors cited and commented on other participants’ arguments, and how often the jurors used logic and reasoning in making their arguments.

For each utterance, we coded whether it included evidence that supported the juror’s argument. Each juror made an average of 4.71 utterances that included evidence, and the highest number of such utterances with evidence for a single juror was 28 (Table 4). We also coded if the utterance included commenting on another participant’s argument. For example, an utterance is coded as commenting on other participants if the

utterance specifically supported, opposed, or added a supplementary explanation after citing arguments of the judge, prosecutor, lawyer, witness, or other jurors. Jurors made on average 1.86 utterances that commented on other participants, and the highest number of comments on other participants made by a single juror was 8. Finally, we also coded if each utterance used logic or reasoning to support the juror's argument. The jurors made an average of 3.18 remarks with logic and reasoning, and the highest number of remarks with logic and reasoning made by a single juror was 16. The quality of deliberations was determined by adding these three kinds of meaningful utterances made by each juror. The jurors averaged 6.34 quality remarks during deliberation. There were jurors who did not make any of these utterances throughout deliberations. The highest number of quality utterances made by a single juror was 52.

Table 4: Values of the Individual Level Variables

	N	Minimum	Maximum	Average	SD
Evidence	242	0	28	4.71	4.261
Commenting on other participants	242	0	8	1.86	1.365
Logic and reasoning	242	0	16	3.18	2.434
Quality of deliberation	242	0	52	6.34	6.467

V. RESULTS

A. Gender and Age Diversity and Quality of Deliberations

First, we examined the quality of deliberations, demographic diversity, and speech diversity between venire juries and recruited juries. While the number of utterances did not vary, the quality of deliberations differed between the two groups (Table 5). The average values of the quality of deliberations were 4.76 in venire juries and 7.62 in recruited juries ($t = -3.715$, $df = 211$, $p = 0.000$).⁶⁶ The fact that the quality of deliberations in recruited juries is higher is arguably because these jurors were recruited through a voluntary application process through the homepage of the Supreme Court. Respondents tended to be active citizens who are interested in the jury trial system, as well as law students, who began the process with more knowledge of the trial process. Demographic diversity was lower in the recruited jury group (0.74 compared to 1.08), likely for the same reason ($t = -10.802$, $df = 242$, $p = 0.037$). This clear distinction between the characteristics of the venire juries and the recruited juries suggests a need for us to consider these two groups separately in examining the effect of

⁶⁶ We used ANOVA (analysis of variance) for these statistical analyses.

diversity on the quality of deliberations. Speech diversity did not differ between the venire jury and the recruited jury groups ($t=2.302$, $df=240$, $p=0.519$).

Table 5: Values of the Jury-Level Variables of Different Groups

Variable	Group	N	Mean	SD
Quality of deliberations***	Venire jury	108	4.76	4.10
	Recruited jury	134	7.62	7.65
Demographic diversity*	Venire jury	108	1.08	0.21
	Recruited jury	134	0.74	0.26
Speech diversity	Venire jury	108	0.92	0.04
	Recruited jury	134	0.92	0.07

*** $p < 0.001$, * $p < 0.05$.

B. *Correlation of the Variables*

We examined the correlation of the variables and found that the type of jury group (venire or recruited) is correlated with demographic diversity, speech diversity, and the quality of deliberations (Table 6). Recruited juries had less demographic diversity and higher quality of deliberations. Demographic diversity and speech diversity was also positively correlated in that higher demographic diversity corresponded to higher speech diversity. Speech diversity was also positively correlated with argument engagement and quality of deliberations.

Table 6: Correlation of Variables

	Jury group	Demographic diversity	Speech diversity	Argument engagement	Quality of deliberations
Jury group	1	-0.572**	-0.051	-0.080	0.220**
Demographic diversity	-0.572**	1	0.206**	-0.122	-0.079
Speech diversity	-0.051	0.206**	1	0.222**	0.385**
Argument engagement	-0.080	-0.122	0.222**	1	0.076
Quality of deliberations	0.220**	-0.079	0.385**	0.076	1

** $p < 0.01$

C. Diversity and Quality of Deliberations

We examined how demographic diversity and speech diversity influence the quality of deliberations. We found that both of them influence the quality of deliberations ($F=11.54$, $p=0.001$, and $F=56.26$, $p=0.000$, respectively), and that the correlation between the two variables also influences the quality of deliberations ($F=6.89$, $p=0.009$; Table 7). When the level of demographic diversity is low, the quality of deliberations was higher than when the level of demographic diversity is high, regardless of the level of speech diversity of the group. Contrary to literature in the United States focusing on the benefits of diversity within American juries, the results of our study in Korea demonstrated that the level of demographic (age and gender) diversity and the quality of deliberations were negatively correlated. However, speech diversity was positively correlated with the quality of deliberations. The higher the level of speech diversity, the higher the quality of deliberations was. The quality of deliberations was at the highest level when demographic diversity was low and speech diversity was high (Figure 1). This result may be driven by the fact that the recruited jury group, consisting of volunteer applicants, was less diverse but more knowledgeable and interested in the jury trial process. Therefore, we decided to examine the two groups separately in conducting our analyses of the relationship between diversity and the quality of deliberations.

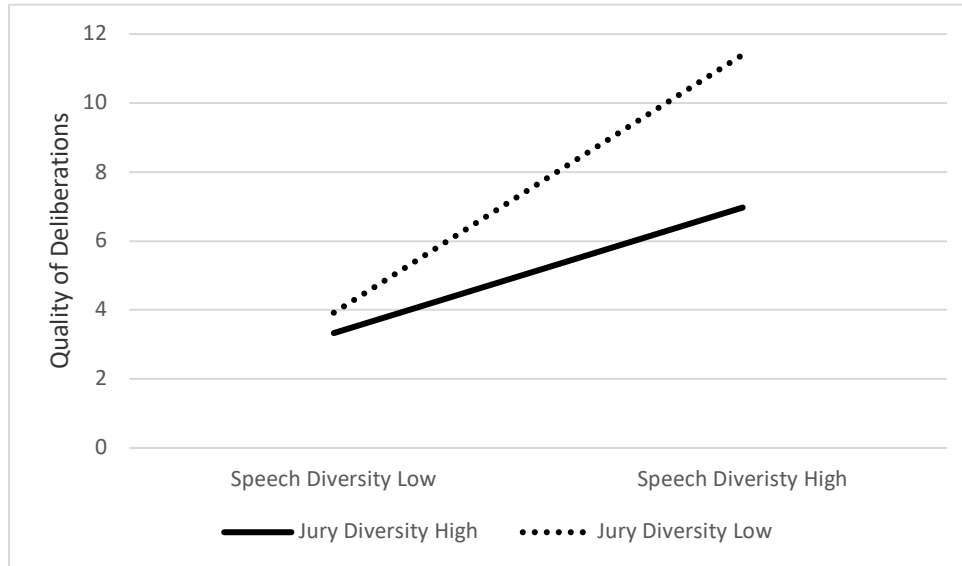


Figure 1: Quality of deliberations by demographic diversity and speech diversity

Table 7: Quality of Deliberations by Demographic Diversity and Speech Diversity

	Sum Sq	df	Mean Sq	F	p-value
Speech diversity	1813	1	1813	56.26***	0.000

Demographic diversity	372	1	372	11.54***	0.001
Speech diversity * dem. diversity interaction	222	1	222	6.89**	0.009
Residuals	7671	238	32		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

D. Demographic Diversity and Quality of Deliberations

Demographic diversity (age and gender) negatively influenced the quality of deliberations ($F=7.37$, $p=0.006$; Table 8). Jury group characteristics (whether the jury was a recruited jury or venire jury) also influenced the quality of deliberations, such that ($F=5.53$, $p=0.020$; Table 8) the quality of deliberations was higher in the recruited jury group than in the venire jury group (Figure 2). There was no interaction effect of demographic diversity and jury characteristics on the quality of deliberations. However, the negative relationship between demographic diversity and the quality of deliberations was found only in the venire jury group (Figure 2).

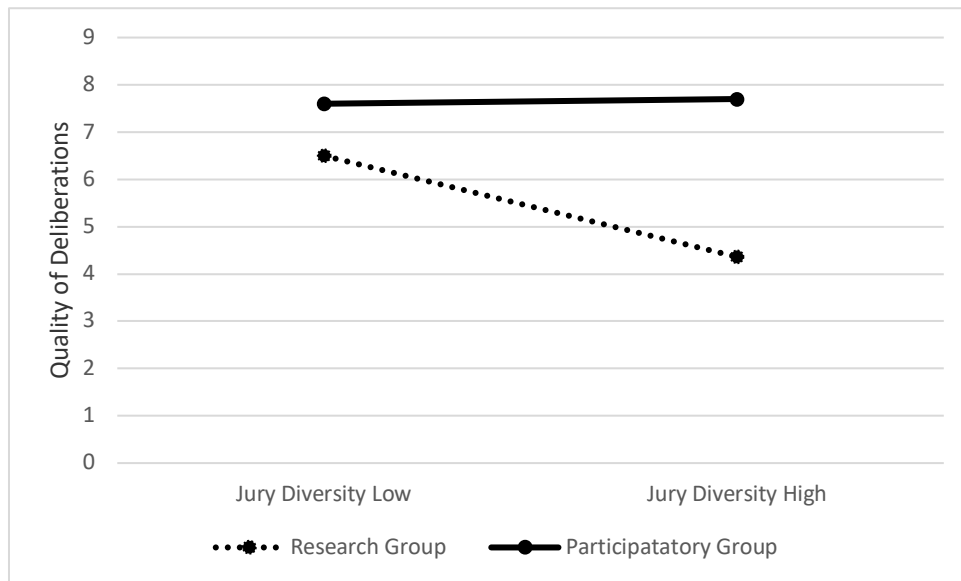


Figure 2: Quality of deliberations by demographic diversity and jury characteristics

Table 8: Quality of Deliberations by Demographic Diversity and Jury Group

	Sum Sq	df	Mean Sq	F	P-value
Demographic diversity	295	1	295	7.37***	0.007
Jury group	221	1	221	5.53*	0.020
Demographic diversity * jury group interaction	48	1	48	1.20	0.274

Residuals 9571.1 238 40.215

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

E. *Speech Diversity and Quality of Deliberations*

Both speech diversity and the characteristics of the jury (recruited or venire jury) influenced the quality of deliberations ($F=56.60$, $p=0.000$ and $F=8.53$, $p=0.004$ respectively; Table 9). When jurors spoke more evenly, as opposed to one or two jurors dominating the discussion, the jurors tended to engage in better deliberations by more often providing evidence, commenting on other participants' arguments and using logic and reasoning. In addition, the interaction of speech diversity and the characteristics of the jury influenced the quality of deliberations ($F=11.45$, $p=0.001$; Table 9), such that the quality of deliberations was highest for juries that consisted of recruited jurors and are engaged in a high level of speech diversity (Figure 3). Jurors in a jury with more speech diversity engaged in better deliberations in terms of providing evidence, commenting on other participants' arguments and using logic and reasoning in making their arguments.

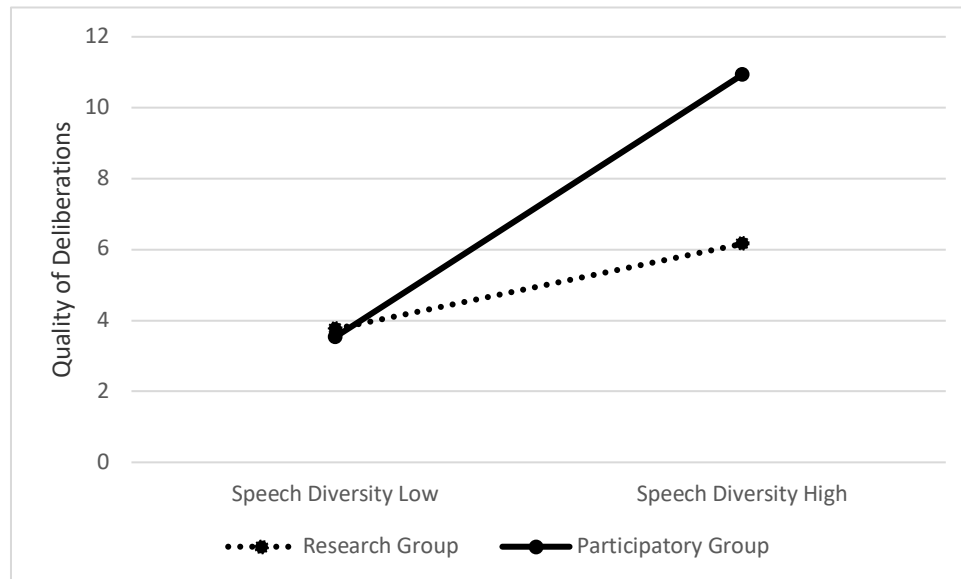


Figure 3: Quality of deliberations by speech diversity and jury characteristics

Table 9: Quality of Deliberations by Speech Diversity and Jury Group

	Sum Sq	df	Mean Sq	F	p-value
Speech diversity	1813	1	1813	56.60**	0.004
Jury group	273	1	273	8.53***	0.000

Speech diversity*jury group interaction	367	1	367	11.45***	0.001
Residuals	7625	238	32		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

F. *Interaction Effect of Argument Engagement*

When we examined how argument engagement and speech diversity influenced the quality of deliberations, we found that higher speech diversity correlated with higher quality of deliberations ($F=54.65$, $p=0.000$; Table 10). Argument engagement itself was not found to influence the quality of deliberations. However, the interaction effect of argument engagement and speech diversity was striking ($F=8.33$, $p=0.000$; Table 10). When speech diversity was high, the quality of deliberations was highest when argument engagement was mid-level (Figure 4). When there was no argument engagement regarding each issue, the quality of deliberations was at the lowest level. At the other extreme, when there was a strong level of engagement, such as more than three separate speech interactions on the issues, the quality of deliberations was also low. In contrast, the quality of deliberations was at its highest level when the speech interactions were exchanged between one and three times. This finding indicates that when jurors commented evenly among themselves, they tended to provide more evidence, comment more on other participants, and use more logic and reasoning. But these indicia of quality deliberations became less common when arguments became more heated through a high level of argument engagement.

When speech diversity was low, on the other hand, the relationship between argument engagement and the quality of deliberations followed a generally opposite pattern. When there was no engagement, the quality of deliberations was low (Figure 4). But when there was a moderate degree of engagement, the quality of deliberations became even lower. Quality of deliberations was highest when there was strong level of engagement. When the jurors did not make remarks evenly, only strong engagement through extended exchanges was able to induce the jurors to provide more evidence, comment on other participants, and use more logic and reason. This sharply contrasted with the finding that, when jurors spoke more evenly, moderate engagement enhanced the quality of deliberations, while strong engagement reduced the quality of deliberations.

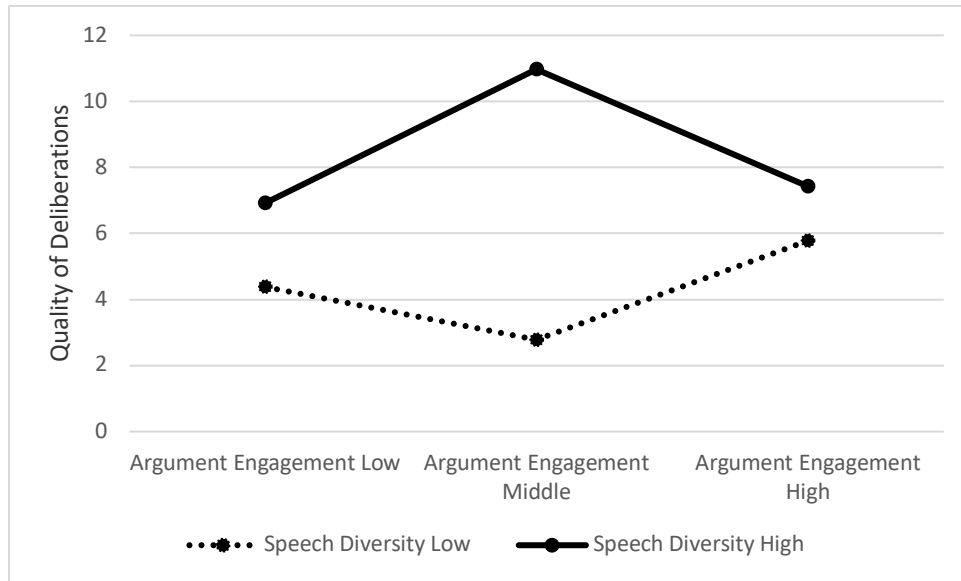


Figure 4: Quality of deliberations by speech diversity and argument engagement

Table 10: Quality of Deliberations by Speech Diversity and Argument Engagement

	Sum Sq	df	Mean Sq	F	p-value
Argument engagement	63	2	32	0.97	0.381
Speech diversity	1781	1	1781	54.65***	0.000
Argument engagement*speech diversity interaction	543	2	272	8.33***	0.000
Residuals	7691	236	33		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

VI. DISCUSSION AND CONCLUSION

In our study of a rich and unique data set of Korean jury deliberations, we focused on examining the dynamics of jury deliberations and attempted to enhance the understanding of a neglected aspect of decision-making: the deliberation process. We were interested in whether and how both the structure of jury composition and the dynamics of the deliberation processes influenced deliberation quality. To that end, we examined how demographic diversity and speech diversity influenced the quality of deliberations. Our findings demonstrated that gender and age diversity negatively influenced the quality of deliberations. These findings contrasted with studies of American juries that have demonstrated the benefits of diversity.

This result suggests that, in Korea, there may be factors that inhibit quality of deliberations in juries that consist of different genders and diverse age groups. Although this was not studied directly, informal observation of

deliberations revealed possible reasons that might prevent quality deliberations. When a jury consisted of some male and some female jurors, one or two of the male jurors often dominated the discussion, to the preclusion of participation by other jurors. This male dominance during deliberations seemed to occur more when the female jurors were younger than the dominant male jurors, but less frequently when the jury consisted of mostly male jurors. Future research is needed to reach a definitive conclusion regarding the occurrence of this male dominance and how it affects the quality of deliberations.

More interesting is the finding that this negative relationship between gender and age diversity and quality of deliberations did not occur in recruited juries where jurors were younger, more educated, and more interested in the jury system. In a jury of more knowledgeable and interested jurors, minimal diversity did not generate a low level of deliberation quality. It is difficult to suggest direct policy implications from this finding, because juries cannot (and should not) be engineered to consist only of younger, more educated and more interested people. Facing a rather contrary demand to promote quality deliberations of juries consisting of members from a broad spectrum of society, we suggest that further research into the dynamics of the deliberation processes is needed to seek an explanation for why demographic diversity produces a contrary result (lower-quality deliberations) in juries with certain characteristics, but not in the case of juries with other characteristics. Future research should evaluate the possible role of the jury moderator, and in particular measuring how the moderator's behavior influenced the dynamics of the deliberation process.

We also found that speech diversity is an important factor that promotes better deliberations. When jurors were able to talk more evenly among themselves, they provided more evidence in making arguments, commented more frequently on other participants' arguments, and used more logic and reasoning in making their arguments. On the other hand, when one or two jurors dominated the discussion process, the jurors altogether in the group tended to provide less evidence, use less logic and reasoning and comment less on the views of other participants. This positive relationship between speech diversity and the quality of deliberations sustained across different jury groups with different characteristics. This finding suggests that more attention should be given to the dynamics of deliberation processes of Korean juries, because age and gender diversity is not sufficient in itself to achieve better deliberations, and may even hinder quality deliberations. We found that when jurors do not take turns and do not provide opportunities for each juror to express thoughts and opinions, jury diversity does not necessarily lead to a more meaningful deliberation with more evidence, reasoning, and responsive comments. The dynamics of the deliberations, not the age and gender diversity of the jury, seems to be the key for quality deliberations.

Speech diversity also promoted high-quality deliberations in relation to the level of argument engagement. When jurors made remarks more evenly among themselves, a moderate level of engagement promoted quality deliberations. When certain jurors made asymmetrically more remarks than others, however, very strong levels of engagement (such as more than three commenting exchanges on each issue) were needed for jurors to provide more evidence, logic and reasoning and to comment on other participants' remarks. Thus, quality deliberations are not only associated with speech diversity but also with argument interactions. Further study is needed to explore how the level of active engagement during deliberations influences their quality, but this finding provides us with another reason to focus on the ways in which Korean jury deliberations unfold, not on who the jurors are.

The results of this study present important implications for a country where a jury system is not an embedded historical asset but an imported system. Especially in a country that has different cultural underpinnings from the United States and other western countries, cultural factors should be given greater consideration, and it is possible that more active intervention will be needed to educate jurors. The results of our study seem to reflect the cultural tendency of Korea – or perhaps East Asian countries more generally – for fluid conversation to be difficult in mixed gender and mixed age groups. Younger people are discouraged to speak out in front of older people, and oftentimes conversation is dominated by the oldest person in the group. This phenomenon is compounded if the group consists of both male and female members. Women tend to be discouraged from being outspoken in general, especially in front of people with whom they are not familiar, and younger women seem to feel this pressure to an even greater extent. Therefore, courts may be able to promote better deliberations by designing the juror's guide and other instruction materials with this cultural factor in mind and explaining the importance of balanced participation to the jurors. The role of the jury moderator may also be critical in this regard.

Our study also provides more general implications for the jury system. Previous literature on American jury diversity has focused its policy recommendations on the jury selection process, but jury diversity seems to interact with other factors that also influence the deliberation process. It would be worthwhile in the Korean context to consider different procedural elements that could encourage each juror to be more evenly and actively engaged in the deliberations, thus promoting more meaningful arguments and conclusions through the deliberation process. Considering the moderator selection process more systematically and improving the juror's guide are examples of such elements, and additional targeted research on the jury deliberation process in more detail would provide further insight as to how jury deliberations can be of higher quality.

Another contribution of our study may come from the heretofore rare methodology of videotaping actual deliberations of shadow jurors who

were potential jurors from the venire (rather than the college students often recruited in mock jury studies), and who participated in the actual trial process in the same way as actual jurors. With a research design devised to imitate actual jury deliberations as much as possible, we revealed that juries with more diverse age and gender jurors do not necessarily engage in higher-quality deliberations, but juries in which jurors speak more evenly among themselves do produce a higher quality of deliberations.