

令和元年6月21日現在

機関番号：34315

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K21491

研究課題名(和文) 裁判員裁判評議を想定した集団討議実験と大型模擬裁判による比較の試み

研究課題名(英文) Comparison between experimental group decision making and mock deliberation in mixed jury trial.

研究代表者

若林 宏輔 (Wakabayashi, Kosuke)

立命館大学・総合心理学部・准教授

研究者番号：40707783

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究は(1)模擬歳馬人裁判・評議データの収集(1-a)と分析(1-b)を並行させながら、(2)小集団評議実験と(3)模擬裁判員裁判実験の比較を目的としていた。しかしながら、目的(2)(3)は、裁判官役の妥当性を実験的に再現することが困難であったため計画を中止した。結果、(1)に関しては評議発言データを分析する手法について大幅な進展が得られた。分析手法の精緻化が図られ、評議内の構成員の関心概念と立場間の構造と、その時間的变化を捉える分析手法を開発した。また当初の研究計画の代わりに、市民の量刑判断への裁判官の影響力を推定する実験を行い、評議内の量刑相場提示の影響を明らかにした。

研究成果の学術的意義や社会的意義

申請時および本研究開始当初は、2009年5月の裁判員制度開始から6年が経過しており、また2019年には制度開始10周年を迎えることが予測されていた。裁判員裁判は日本司法の民主制の象徴であり、中でも市民が裁判員として参加する評議による意思決定は、日本のあらゆる政策決定プロセスの中で最も市民意見を直接的に反映する仕組みである。ゆえに本研究が検討した栽培ん裁判・評議分析技術の開発は、この制度を検証する上で一定の価値を持つ。同手法を用いて今後の制度見直し等の際に情報提供することが可能である。

研究成果の概要(英文)：This study aimed to develop the way of analysis for a deliberation of Japanese mixed jury trial (Saiban in seido) in the criminal justice trial. On our early plan, we tried to compare between some experiments of group decision making and a designed mock trials. But we needed to cancel those which was planned, because we could not prepare the professional judges with high validity in the experiment. We, therefore, collected the protocol data of deliberation which was conducted by legal professionals as mock mixed jury trial. Then we have analyzed the deliberation protocol by the text mining method that we are developing. As a result, our analysis could be developed on three aspects below. 1) to include the time sequences of deliberation for analysis. 2) to find more appropriate statistical analysis, and 3) how should we put the index for the time sequence of deliberation.

研究分野：法心理学

キーワード：法心理学 裁判員制度 評議 可視化

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

申請時および本研究開始当初は、2009年5月の裁判員制度開始から6年が経過しており、また2019年には制度開始10周年を迎えることが予測されていた。裁判員裁判は日本司法の民主制の象徴であり、中でも市民が裁判員として参加する評議による意思決定は、日本のあらゆる政策決定プロセスの中で最も市民意見を直接的に反映する仕組みである。ゆえに評議の在り方について検討する本研究が提案する評議分析技術の開発は重要な意味を持つ。また得られた結果から制度自体の検討や見直しを図る際に情報提供することが期待された。

2. 研究の目的

本研究計画は平成28年度から平成29年度に渡り、まず(1)地方裁判所模擬裁判データの収集(1-a)と分析(1-b)を並行させながら、(2)小集団評議の実験を行った比較、そして総括として最終年度(平成30年度)に(3)模擬裁判員裁判実験を行い蓄積した研究知見を確かめることを目的としていた。しかしながら、目的(2)(3)に関しては、評議内の人数比を操作して複数の模擬評議を行う実験計画であったが、裁判官役の妥当性を実験的に再現することが困難であった。よって2017年度から裁判員評議を更に抽象化し、量刑相場の提示者としての裁判官と裁判官以外(主として意思支援システム等)を比較する心理学実験を行い、(2β)市民の量刑判断への裁判官の影響力を推定する実験を行った。

3. 研究の方法

(1-a) 地方裁判所模擬裁判データの収集

2009年の裁判員制度開始までに各地方裁判所および法曹三者で行われた模擬裁判員裁判の評議データ、または以降も各地方裁判所で不定期に行われる模擬裁判員裁判のデータの収集を行った。各地弁護士会等に協力を仰ぎ、模擬裁判等が行われる際に同評議の記録等を提供頂いた。

(1-b) 地方裁判所模擬裁判データの分析

【分析対象】某地方裁判所で行われた模擬裁判員裁判・評議プロトコルであった。(1-a)で収集したもののうち、1)データに欠損がない(評議だけで公判記録がない。評議の会話が不完全等)、2)法曹三者(裁判所、検察庁、弁護士会)の協力で行われており、公判・評議ともに現職の裁判官3名と市民6名によって行われたものを対象とした。同評議データの裁判員6名は、実際の手続きと同じ裁判員候補者名簿から選ばれた。公判も裁判官、検察官および弁護士が参加した。被告人および公判中に登場する証人等は俳優に依頼されていた。

以下では、主として上記条件を満たしていた評議(計2)のうち1つを主に対象として分析手法の検討・開発を行った。主たる分析手順は以下であった。

①**模擬評議内発話の逐語録化** 上記模擬裁判評議は全て録音・録画されており、評議内の発話は全て逐語録化して分析対象とした。逐語録は発話者の1ターン毎(他者からの発言が挟まれず続いた発話)に対し、評議参加者を匿名化するため、裁判官3名(C, T, S)と裁判員6名(1L-6L)とタグ付して逐語録化した。

②**同逐語録へのテキストマイニング** ①の発話者と発話内容が紐付けられた逐語録に対しテキストマイニングを行った。ソフトウェアはTiny Text Miner (TTM, v.0.89, McCab)を使用し、分かち書き分析から形態素(名詞、形容詞)の出現頻度を算出した。またこれらの形態素のうち、(1)不要語(「この」「あの」等の指示代名詞、また単語だけでは意味が理解できない「こと」「とき」等)を削除し、(2)使用頻度の上位50語を分析対象とした。分析語数に関しては50語以上も利用可能であるが、(1)語数が増えると結果の図の視認性が悪くなる、(2)低頻度語が増えても図の構造に大きな変化がないことから50語を妥当とし、以下の分析を行った。

③発話者×高頻度語(上位50語)の行列に対するコレスポネンス分析

②のテキストマイニングの結果得られた行列(高頻度語50語×評議構成員(9名:裁判官3名、裁判員6名))に対してコレスポネンス分析(以下CA)を行った。コレスポネンス分析は、各行列の要素(各高頻度語と各評議構成員)の類似度(距離)と、その類似度を評価する際の次元(軸)を行・列いずれかのn-1次元分抽出する。このうち主にデータ説明率(累積寄与率)の高い主要な2次元を用いて、各行列要素を同2次元空間上に布置することができる。同2次元の類似度から各構成員と、彼らがよく用いた概念の関係を可視化する。

(2β) 市民の量刑判断への裁判官の影響力を推定する実験

裁判員裁判・評議において、事実認定を争う事件よりも、事実認定に争いはなく有罪は確定しているが、その量刑判断が評議の中心的議題となる事例がある。最高裁判所事務総局(2018)によれば、平成29年度裁判員裁判終局総数900のうち自白事件417件、否認事件は483であった。自白事件では、基本的に、自白の任意性に争いが無い場合は、被告人が罪を認めており有罪判決はほぼ決定的であり、評議ではその被告人にどの程度の刑を課すかが議論となる。このとき、

一般的に裁判官は評議内で、過去の同等事件の量刑判断をグラフ化した、いわゆる量刑相場を市民に提示することが知られている。しかし裁判員が専門家からの情報として量刑相場を受け取った場合、同情報が市民の意思決定に影響を与える可能性は十分にある。心理学では、このように、人がある問題について考える時に、先行する関連情報を無視することが出来ず、判断時にこれらの情報の影響力を残した判断が行われる傾向を「係留—調整ヒューリスティクス（通称：アンカリング効果）」と呼ぶ。

そこで本研究（2β）では、評議内で提示される量刑相場が裁判員の量刑判断に与える影響を心理学実験で検討した。また量刑相場の提示が裁判官以外から提示された場合を比較対象として設定した。近年、AI 等による意思決定支援システムの効用が問われており、これは司法においても同様であると考えられる。よって本実験 1 では、裁判官(human)が過去の裁判所内の相場を提示する条件と、意思決定支援システム(System)が算出した相場が提示される条件の比較を行った。ま相場の妥当性の高低(窃盗障害事件での相場 4 年と相場 10 年)も比較条件とした。

4. 研究成果

(1-a)評議データの収集

評議データの収集については、研究期間中に、各地方裁判所および弁護士会等で行われた模擬裁判員裁判のデータを計 17 評議分収集することができた。ただし、そのうち上記方法で示した条件を満たした評議は 2 つだけであった。各地方裁判所で行われる模擬裁判員裁判のうち法曹三者が揃って行われるケースは珍しく、現職の裁判官が模擬評議でどのように振る舞うかがわかる貴重なデータである。一方で、法曹三者で行われる模擬裁判員裁判は、三者による刑事訴訟法手続きの確認的側面が強く、実験データとしての認識は薄い。また分析手法がないために、記録はされるものの、その逐語録化や分析に向けた公開といった視点が欠けている。同データは参審制裁判をとる全ての国にとって貴重なデータであり、今後これらのデータの公開や記録化の整備が望ましいと考えられる。

(1-b)評議データの分析方法の開発

①これまでのテキストマイニングによる評議分析手法

上記研究手法で提示した分析手順に従って得た、評議の高頻度語 50 語×評議構成員（9 名：裁判官 3 名、裁判員 6 名）の行列に対し、CA を行った結果を以下に示す（図 1）。CA の結果、主要 2 次元の累積寄与率は 43.2% と 35.8%（計 79.0%）であった。

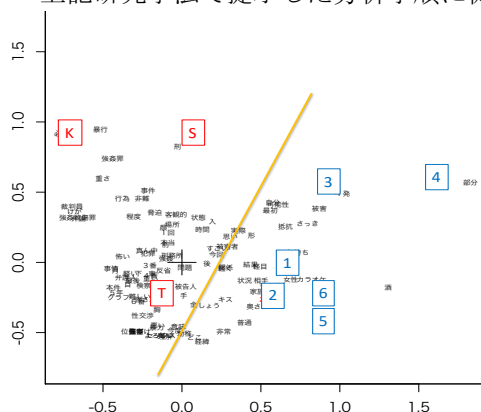


図 1 研究期間中に得られた模擬裁判員裁判の評議のテキストマイニングに基づく布置図：裁判官 (T,S,K)、裁判員 (1~6) と斜線は申請者の強調

各参加者の図中の布置関係をみると裁判官 (T,S,K) は図中左側に、裁判員 (1~6) は図中右側に別れて布置された。これは両者が評議内で高頻度で使用した語（注目した概念）が大きく異なっていたことを意味する。また CA の性質から図中の端に位置する語は、その周囲の話者が主に発言した特徴的な話題として評価できる。つまり裁判員 4 番は「部分」、3 番は「1 発」、また裁判員全体は「酒」に、裁判官よりも着目していたことを意味する。

一方、同様に CA の性質から、原点 (0,0) に近い構成員は使用頻度が非常に高いこと、また全体の語を一定数使用した際に布置される。裁判官 T は裁判長であり、分析結果から裁判長が評議内で多様な話題に触れ、また他の裁判官と比べて発言頻度も高いことがわかる。

同分析結果は申請者がこれまでに取り組んできた評議分析手法であり、本研究期間中に得られた限りなく実際の裁判員裁判に近い評議においても、同様に分析することが可能であった。また申請者がこれまでに分析してきた他の評議と同様に、裁判官（特に長）が図中心に布置されること、裁判員が裁判官とは異なる点に注目する傾向にあることが追認された。

②評議の時系列展開の可視化

評議は複数名の議論のプロセスであるから、評議の開始から意思決定までの時間の流れは一つ重要な分析の次元である。申請者のこれまでの分析手法では、この評議の時間を圧縮し、評議内で使用された語と話者との関係性を空間的に表現してきた。よってこの主要な次元である「時間」を同分析に取り込むために、評議プロトコルを話題の転換毎に質的に分割（フェーズ化）し、分析することを試みた。

まず評議逐語録を申請者が読み込み、話題の転換（フェーズ）と考えられる箇所でテキストを区切った。結果、同評議では16フェーズが得られた。次に同フェーズ毎に、参加者の各発言を分類した。分類タグとして第1フェーズの裁判員1番の発言を「1L1」、第10フェーズの裁判員4番の発言を「10L4」とタグ付した。よって裁判官Kの第16フェーズの発言は「K16」であった。各フェーズにおける各話者を分類したところ計79タグ（各フェーズでの発言がない、および上位50語を使用していないものは削除）となった。

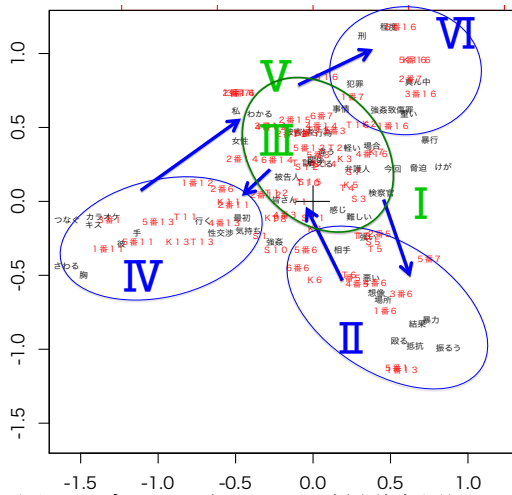


図2 時系列を組み込んだ評議布置図
(図中のI～VI、楕円等は申請者の強調)

図中の話者と高頻度語のポジショニングをまとめると、評議中の話題の展開としては大きく6つの構造が示された(図中I～VI)。まず話題I、III、Vはほぼ図中央に位置しており、各語から話題の内容は特定できないが、一方話題II、IVと時系列的に入れ子となっていることがわかる。よって評議では、頻度が高く関心の高い話題(I・III・V)が繰り返し議論されるが、適宜特定のトピック(II:性的行為直前の暴力と抵抗の程度、IV:カラオケ内での両者の関係)について議論が行われ、改めてI・III・IVの話題に立ち戻るといった評議の時間的展開があったことがわかる。またVIは「判決と量刑」に関する語がまとまり、評議という事例特有の最終段階の話題も同分析から抽出できた。

③使用する統計手法の検討

申請者はこれまで、上記のように評議逐語録にテキストマイニングを行い、得られた行列に対してCAを行う手法を用いてきた。対して多次元尺度構成法(以下、MDS)もポジショニング分析、つまり分析対象の位置関係を表すことを主な目的とする手法である。CAもMDSも、行列(ここでは評議参加者(市民6名と裁判官3名)と高頻度語)から行列各要素の類似度を算出し、その距離を図中で表現する。いずれも評議を解釈するための次元を評議内容から抽出する点で共通しているが、CAは主要2次元を使用するのに対し、MDSは得られた複数の次元を全て図中に表現する点異なる。またこれら各統計分析後の布置図の解釈はこれら抽出された次元に限定される。よって、これら統計手法の違いによる分析結果の比較を行った。

以下に、79タグ×高頻出語50語(名詞、形容詞を抽出)の行列にMDS(metaMDS, Bray-Curtis法, 試行数=20, Stress =0.28, 5次元)またCA(corresp, nf=8, 2次元寄与率39.3%)を行った結果をそれぞれ図3および図4に示す。

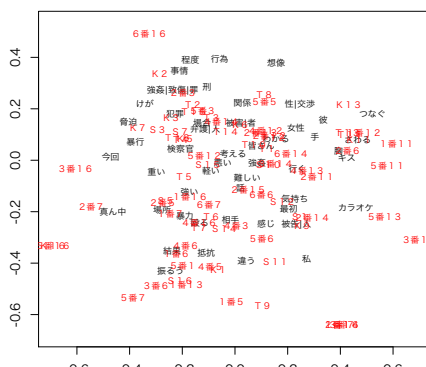


図3 MDSによる布

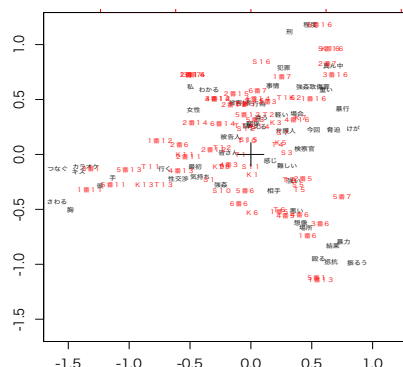


図4 CAによる布置図

まず図3のMDSからは特別なパターンは観察されなかった。

一方で、図4のCAの結果(図2の結果と同じ)からは、先述したように評議内時間による話題の展開が視覚的にも抽出された。

④テキスト分割方法の検討

また上記の分析では、申請者が質的に評議逐語録を読み込み、話題の転換を特定してフェーズ化し、そのタグを CA 分析にかけて布置図を作成したものであった。

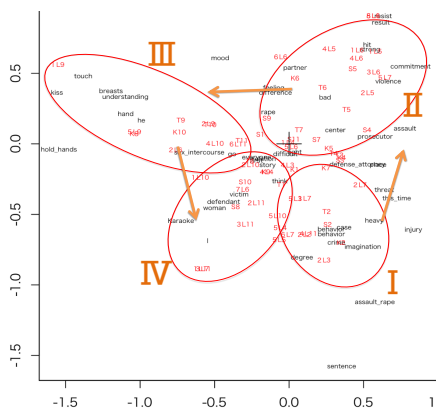


図5 自動テキスト分割による布置図

この意味において、テキストの分割はある意味で、恣意的であり、同手続を自動化しても同様の結果が得られるかを確認した。本分析では、逐語録で話者が60回切り替わるごと自動的にフェーズ化しタグ付けを行った。結果、評議全体で15フェーズとなり、各フェーズにおける話者の発言（タグ）は67得られた。よって67タグ×50語の行列にCAを行った。

結果、得られた布置を図5に示す。図3と比較した場合よりも単純な話題の展開構造が得られた。しかしながら、質的なテキスト分割の結果（図3）と比較して、裁判官-裁判員それぞれの話題（高頻度語）の関係に明確な違いは観察されなかった。

結果的に自動的なテキスト分割（今回は60ターン）は、時間的に近接しているとはいえ、実際の評議内話題の連続性を分断してタグ化してしまうため、分析結果として評議内容の直感的な解釈が反映されない可能性が考えられた。ただし、この分割基準をより細かくすることなどで対応は可能になる可能性が考えられた。

以上が、本申請研究期間中に、裁判員裁判の評議分析手法を検討した成果である。同手法はより現実の裁判員裁判評議に限りなく近いデータに対しても同様の分析結果を得ることができ

る。また新たに開発した時系列的変遷をとり入れた分析手法（図3）により、評議内主要話題と立場の構造だけでなく、その時間的変化も可視化し分析することが可能となった。分析手法としては今後も検討の余地があるが、一事例的な複数名の発話データを分析する手法として、一定の成果を得たと考えられる。

(2β) 市民の量刑判断への裁判官の影響力

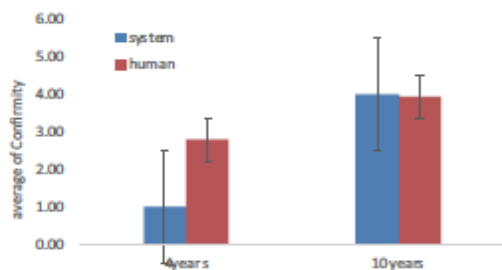


図5 裁判官と意思決定支援システムの量刑相場提示による市民判断への影響

実験1では、提示された相場と市民の量刑判断の差（年）を従属変数とし、二要因分散分析を行ったところ、妥当性の高低の主効果のみが有意($F(1, 45) = 31.297, p = .000$)であり、交互作用が有意傾向($F(1, 45) = 3.160, p = .08$)であった。

結果から、AIを想定した意思決定支援システムが量刑相場を提供したところ、市民の量刑判断の標準偏差（分散）が大きくなった。これは提示された相場とは大きくことなる判断をした市民と、提示された相場を支持した市民に大きく別れたことを意味する。一方で、裁判官（Human）

条件で妥当な量刑（4年）が提示された場合に分散は小さく、市民が平均3年程度は相場と異なる判断を行う傾向が確認された。

次に実験2では、このシステム側の判断が、人間以外のシステムへの不信感や恐怖感といった態度（ロボフォビア Robophobia）から予測できるかについて検討した。NARS (Negative Attitudes toward Robots Scale)の下位項目（S1：ロボットに対するネガティブな相互作用経験、S2：ロボットに対するネガティブな社会的影響、S3：ロボットに対するネガティブ感情）を測定する14個の質問の回答傾向と量刑判断についてPearsonの相関係数を算出した。結果、S1、S2、S3はそれぞれ0.058、0.356、そして0.386となり、有意な相関はなかった。しかし、S2とS3はS1と比べて相対的に量刑判断に影響を与えていた可能性がある。実験1と2の結果より、裁判員裁判評議内における量刑相場の提示については、先行研究と同様にアンカリング効果が認められた。

5. 主な発表論文等 〔雑誌論文〕（計7件）

1. 若林宏輔(2019). 市民の犯罪不安とリスク認知の心理学, 季刊刑事弁護, 98, 138-142 (査読なし)
2. 若林宏輔(2019). 市民と犯罪に関する認識の心理学, 季刊刑事弁護, 97, 203-207 (査読なし)
3. 荒川歩, 滑田明暢, 綿村英一郎, 若林宏輔, 笹倉香奈, 福島由衣 Law and Human Behavior にみる近年の法と心理学研究の動向 法と心理, 7, 70-76. (査読なし)
4. 若林宏輔・林勇吾(2017). 意思決定支援システムは裁判において有効な指標になりうるか? 電子情報通信学会技術研究報告, 117, 55-58. (査読あり)
5. 小坂祐貴・山崎優子・石崎千景・中田友貴・若林宏輔・サトウタツヤ(2016). 裁判員裁判における評議パターンの提案-質的・量的分析の統合から- 立命館大学人間科学研究, 35, 49-67. (査読あり)
6. 若林宏輔(2016). 評議に声は届いているか? 模擬裁判員裁判の公判と評議の心理学的分析, 季刊刑事弁護, 90, 124-133. (査読なし)

[学会発表] (計 13 件)

1. Yugo Hayashi, Kosuke Wakabayashi, Shigen Shimojo & Yukoh Kida (2019). Using Decision Support Systems for Juries in Court: Comparing the Use of Real and CG robots, Human Robot Interaction 2019(HRI2019), Daegu, ROK.,
2. Kosuke Wakabayashi(2018). Visualized deliberation: Analyzing by text mining method through the time sequence Annual Conference of the European association of Psychology and law 2018, Turku, Finland
3. 若林宏輔(2018). 「いつも一人いる」人は評議に現れ、そして存在し続けるのか? 法と心理学会第 19 回大会シンポジウム『映画で考える裁判員制度と心理学』, 法と心理学会第 19 回大会, 関西国際大学
4. Yugo Hayashi, Kosuke Wakabayashi (2017). Influence of Robophobia on Decision Making in a Court Scenario: A Preliminary Experimental Investigation using a Simple Jury Task, Human Robot Interaction 2018, Chicago, IL, U.S.
5. Kosuke Wakabayashi (2017). Inside deliberation: How can we express the sequence of deliberation, 11th East Asian Law and Psychology conference 2018, Taipei National University, Taipei, Taiwan.
6. 若林宏輔 (2017). 評議の可視化: 評議内の時間的変遷と構造, 法と心理学会第 18 回大会, 成城大学,
7. 若林宏輔(2017). 議論の中で人はどのような変化を見せるのか —模擬裁判員裁判・評議のテキスト分析から—, パーソナリティ心理学会第 26 回大会, 東北文教大学
8. 若林宏輔・林勇吾(2017). 意思決定支援システムは裁判において有効な指標になりうるか? 電子情報通信学会, 沖縄産業センター
9. 若林宏輔(2017). 評議の心理学的分析に基づく弁護実践の効果 第 59 回(平成 28 年度) 弁護士夏期講習(近畿地区)評議を見据えた弁護活動はどうあるべきか(招待講演), 近畿弁護士連合会(大阪弁護士連合会)
10. Yugo Hayashi, Kosuke Wakabayashi(2016). Can AI become reliable source to support human decision making in a court scene? The 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing, Portland, OR, U.S.
11. Kosuke Wakabayashi (2016). Inside deliberation: How to get away with not-guilty of murder, 10th East Asian Law and Psychology conference, Chunback university, South Korea.
12. Kosuke Wakabayashi (2016). Analyzing the deliberation process in the mixed jury trial in Japan, from the collaborative aspect in the fairness achievement. the 31st International Congress of Psychology, Yokohama, JAPAN
13. Kosuke Wakabayashi (2016). Why are the numbers of citizens those? Through the comparing of the resistibility of the citizens' number in the mixed jury deliberation. International workshop: Commonsense Justice? Appraising Lay Judge/Jury Criminal Trials, International Institute for the Sociology of Law, Onati, Spain

[図書] (計 1 件)

1. 若林宏輔(2017). 評議を見据えた弁護活動はどうあるべきか—各地の模擬評議の心理学的分析を踏まえて— 日本弁護士連合会(編)日弁連研修叢書 現代法律実務の諸問題<平成 28 年度研修版>, 469-509 (分担執筆)

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。