

Engel, Christoph

Working Paper

Judicial decision-making a survey of the experimental evidence

Discussion Papers of the Max Planck Institute for Research on Collective Goods, No. 2022/6

Provided in Cooperation with:

Max Planck Institute for Research on Collective Goods

Suggested Citation: Engel, Christoph (2022) : Judicial decision-making a survey of the experimental evidence, Discussion Papers of the Max Planck Institute for Research on Collective Goods, No. 2022/6, Max Planck Institute for Research on Collective Goods, Bonn, <https://hdl.handle.net/21.11116/0000-000A-E4EF-8>

This Version is available at:

<https://hdl.handle.net/10419/274055>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Discussion Papers of the
Max Planck Institute for
Research on Collective Goods
2022/6



**Judicial Decision-Making
A Survey of the Experimental
Evidence**

Christoph Engel

MAX PLANCK
SOCIETY





Judicial Decision-Making A Survey of the Experimental Evidence

Christoph Engel

August 2022

Judicial Decision-Making A Survey of the Experimental Evidence

Christoph Engel

Abstract

Judges are human beings. Is their behavior therefore subject to the same effects that psychology and behavioral economics have documented for convenience samples, like university students? Does that fact that they decide on behalf of third parties moderate their behavior? In which ways does the need matter to find a solution when the evidence is inconclusive and contested? How do the multiple institutional safeguards resulting from procedural law, and the ways how the parties use it, affect judicial decision-making? Many of these questions have been put to the experimental test. The paper provides a systematic overview of the rich evidence, points out gaps that still exist, and discusses methodological challenges.

JEL: K10, K13, K14

Keywords: judicial decision-making, bias, heuristic, attitudinal model, ambiguity, parallel constraint satisfaction, public perception

I. Research Questions

Experiments are artificial by design. They remove much, sometimes almost all, of the social context. Strictly speaking, what the experiment studies is only analogue to what the researcher wants to understand. The gap between the experiment and the real world problem that motivates the endeavour is not a bug, it is the definitional feature. Simplifying the situation is a precondition for excluding confounds. If the experiment is properly designed, alternative explanations are ruled out. If there is a difference between the baseline and the treatment, it must result from the manipulation.

But as Albert Einstein has once famously put it: everything should be made as simple as possible – but not simpler. For experimental research on judicial decision-making, the resulting trade-off between internal validity (there are no alternative explanations) and external validity (the experimental result is relevant for the research question) is never easy (Lucas 2003, Schram 2005). Extant research is situated on a continuum. On the internal validity end research starts from the assumption that judicial decision-making is human decision-making. This research hypothesises that knowledge about human decision-making in general is also relevant if these humans happen to be judges. One line of research turns the assumption into a hypothesis, and tests real judges on characteristic features of human decision-making that have been robustly documented with convenience samples. The main contribution of this work is the choice of real judges as a non-standard sample.

A second line of research focuses on differences between the standard tasks that have been used to experimentally isolate the behavioral regularity of interest, and judicial decision-making. The most basic deviation is motivational. Judges do not decide on behalf of themselves. They decide on behalf of the parties. This feature of the task can be isolated without the necessity of recruiting real judges (Cerrone and Engel 2019).

Actually the judicial task is heavily institutionally contained. Judicial decision-making is tightly controlled by a rich set of procedural rules. For the most part, judges are also not free how to decide on substance. They have to expressly justify deviations from the existing doctrine. In the common law jurisdictions, they are additionally bound by precedent. Potentially each of these procedural or substantive constraints has a discernible behavioural effect, as do any of their combinations. If a researcher has reason to believe that an individual constraint has a measurable and sizeable behavioural effect, an experiment can be designed to test the hypothesis. The cumulative effect of multiple constraints is harder to test. In principle, factorial designs are able to test such interactions between impulses. But the approach quickly runs into combinatorial explosion. The researcher needs theory, or at least a good guess, to select the combinations of interest.

An alternative research strategy does not start from behavioural theory, but from the constraint. Such designs compare the otherwise constant setting, once without and once with the constraint. If the researcher wants to maintain as much experimental control as possible, she will use a hypothetical judicial choice. This makes it possible to tailor the remaining features of the task to the specific intervention of interest, either giving the expected effect its best chance, or putting the purported effect to a particularly hard test. If some rule making authority

has a serious interest in the evaluation of the intervention, she may collaborate with a researcher and implement a randomized control trial. Then the intervention is tested under real-life conditions. Yet this biggest advantage of the approach is also a major challenge. The authority may, for instance, prematurely lose interest. Then the study may have to be discontinued, possibly even before enough data has been collected to make the analysis meaningful. Moreover, the room for manoeuvre tends to be limited. The design that may be most appealing from an academic perspective, for instance for isolating behavioral channels, may appear too disturbing to the court administration.

Judicial decision-making is a particularly challenging object of study. At court, the parties' well-being is at stake. Just assigning otherwise sufficiently similar cases to manipulated conditions may not be ethically or politically feasible. In the worst of all cases, from the perspective of the empirical researcher, a true control group is out of the question. Then at most differently designed interventions may be compared.

Procedural codes worry about impartiality, and about the perception of impartiality in the population. One safeguard of impartiality is confidentiality. This safeguard becomes particularly relevant if one accepts that judges effectively have leeway, even in the continental jurisdictions. Were to be known which element of procedure has been most important for the disposition of the case, there would be more scope for forum shopping. This concern is imminent if the judicial decision is taken by a panel, a bench of professional judges, or the jury. To contain this risk, the law makes direct observation, let alone manipulation, impossible. Behavioural research must find indirect paths. The classic reaction is studying mock juries, i.e. panels of jury eligible participants on materials that resemble real jury trials (a classic is Hastie, Penrod et al. 1983). This demonstrates a benefit from experimentation that transcends random assignment to treatment. Shifting from real to hypothetical cases, and real to hypothetical juries, makes social interactions observable that otherwise are purposefully concealed.

A further concern rests on a robust empirical observation. Most individuals rarely, if ever, directly interact with the judicial system. How they, or their relatives or close friends, feel treated on these rare occasions is highly predictive for their willingness to obey the law if enforcement is unlikely (Tyler 2006). Hence there is a pronounced spillover to other areas of life. Randomly exposing individuals to experimental conditions that might trigger such a negative spillover could therefore have a negative externality on potential victims, or on society at large. This normatively undesirable effect is a further reason for tweaking the experimental method and shifting from actual to hypothetical cases.

The following survey brackets jury research. There is so much experimental work on the topic that it would require a separate paper (see only the following surveys and meta-studies Gerbasi, Zuckerman et al. 1977, MacCoun 1989, MacCoun 1990, Steblay, Besirevic et al. 1999, Devine, Clayton et al. 2001, Mitchell, Haw et al. 2005, Steblay, Hosch et al. 2006, Desmarais and Don Read 2011, Wiener, Krauss et al. 2011, Bornstein, Golding et al. 2017, Simon 2019, Hoetger, Devine et al. 2022).

II. Experimental Findings

1. Heuristics and Biases

A rich strand of the literature asks to which degree behavioral regularities that have been robustly documented in other populations matter for judicial decision-making (for an overview see Guthrie, Rachlinski et al. 2001, Teichman and Zamir 2014), including specialized panels within their professional area of expertise (Rachlinski, Guthrie et al. 2006).

Judges are subject to anchoring. When faced with a quantitative task, but not having exact numeric information, any number to which they compare the unknown quantity affects their assessment. This effect can be strategically exploited by the parties, by throwing in a number that leads to a reduction or increase of the estimate, as serves their cause (Bystranowski, Janik et al. 2021). Judges are liable to framing. Whether a decision is presented as a gain or as a loss affects how they evaluate it (Guthrie, Rachlinski et al. 2001: 794). The same holds for the way how an expert reports estimates of risk (Scurich and John 2011). Anchoring and framing have been rationalized with salience (Bordalo, Gennaioli et al. 2015).

Asylum judges fall prey to the gambler's fallacy. They underestimate the probability of a sequential streak (Chen, Moskowitz et al. 2016). In experiments, untrained participants are also inconsistent in the way how they translate their qualitative assessment into dollar amounts, for instance about punitive damages; yet the effect does not translate to observational data (Eisenberg, Rachlinski et al. 2001).

Judges exhibit hindsight bias. When learning that a possible event has actually happened, they increase the estimated probability that it would have happened (Anderson, Jennings et al. 1997, Guthrie, Rachlinski et al. 2001: 799, Oeberst and Goeckenjan 2016). Judges neglect the baseline. If the conditional probability of an event is very high, but the unconditional probability (the baseline) is very low, judges tend to come up with an exaggerated assessment (Guthrie, Rachlinski et al. 2001: 808). When decision-makers face a portfolio of decisions, evaluations spill over. If there is one severe case, the remaining cases are evaluated more favorably (Kelman, Rottenstreich et al. 1996).

Judges have a hard time handling probabilities. This leads to miscarriage of justice. Judges do not acquire distributional information when it is available (Sonnemans and Van Dijk 2012).

Judges do not only infer from the emotions of victims how deserving they are (Rose, Nadler et al. 2006). Emotions do also affect how they evaluate the case (Feigenson and Park 2006, Wistrich, Rachlinski et al. 2014).

2. Personal Attitudes

The foregoing are cognitive limitations. Unless they decide about phenomena that affect society at large, like environmental damage, the case disposition has not direct effect on the judge's personal well-being. Outright selfishness can therefore at most be a minor motive

(Posner 1993), e.g. when preparing for reelection or promotion (Shepherd 2009). The prominent attitudinal model does, however, expect judges to be swayed by ideology (Segal and Spaeth 1993) or, more sophisticatedly, to strategically take into account which decision serves their ideological cause best (Epstein, Knight et al. 2001). Experimental evidence qualifies this expectation. Judges decide on behalf of others. In the lab, participants even serve their experimental constituency if they have no personal stake in it, and if they have to invest a fraction of their experimental endowment. They are motivated to do a good job, even if it has been randomly assigned to them (Engel and Zhurakhovska 2017). In one vignette study, laypersons instructed to decide as judges exhibit ideological bias (Kantorowicz-Reznichenko, Kantorowicz et al. 2022). In another vignette study, a correlation of some, but not all, political opinions with the proposed decision of the case is observed (Braman 2006). In a further study, professional judges, unlike laypersons, are unaffected by their personal attitude towards salient policy issues (Kahan, Hoffman et al. 2015).

Even if decision makers are motivated to be impartial, bias may sneak in inadvertently. This has been demonstrated for race (Rachlinski, Johnson et al. 2008) and ethnicity (Gazal-Ayal and Sulitzeanu-Kenan 2010, Choi, Harris et al. 2022), gender (Kulik, Perry et al. 2003, Miller 2019), ideology (Furgeson, Babcock et al. 2008) or induced punishment philosophy (McFatter 1978), and for the stereotype that tattoos are typical for criminals, and hence predict guilt (Funk and Todorov 2013). Legal education mitigates the effect of stereotypes, but only if the learned rules are simple, rather than judgmental - rules, rather than standards, in the terminology often used in law and economics (Girvan 2016).

In international courts, a further potential for bias results from nationality. If a judge on the bench has the nationality of one of the parties, this may bias the decision in this nation's favor. Random assignment to panels in the appellate body of the World Trade Organization creates quasi-experimental evidence. Overall, there is indeed a bias, but it is driven by the dominance of judges having US nationality (Arias 2019).

3. Ambiguity

In the effects listed in the previous section, the bias had a chance due to the characteristic complexity of legal cases, which gives room for motivated reasoning (Norton, Sommers et al. 2006). The literature on judgement and decision making has characterized the typical task in the courtroom as ambiguous (Simon 2004). Actually, the ambiguity tends to be severe. Judges do not only lack credible information about the base rate of the crime with which the defendant has been charged, and the conditional probability of him being guilty, given the available evidence. In most cases, even the state space is not well defined: multiple chains of events might have happened (more on these distinctions from Engel and Goerg 2018). The judge cannot avoid engaging in sense making (Kunda 1999).

It has been shown (with laypersons) that individuals tend to do so by way of storytelling. They try to combine the pieces of evidence to a consistent story (Hastie, Penrod et al. 1983, Pennington and Hastie 1986, Pennington and Hastie 1991). The underlying mental process has been modelled as parallel constraint satisfaction (Simon 1998). In an iterative process

(Simon, Pham et al. 2001), decision makers reason bidirectionally: from the evidence to the interpretation, but also back from the provisional interpretation to the evidence (Holyoak and Simon 1999), until a coherent interpretation has been forged that meets the probative standard, like preponderance of the evidence or, more strictly, beyond a reasonable doubt (Glöckner and Engel 2013).

This mental mechanism is highly adaptive. It enables decision makers to make reasonable choices although they full well know that their knowledge and understanding are incomplete. But this human capacity comes at a price. Decision makers can become partisan, while subjectively believing that they are impartial (Guthrie, Rachlinski et al. 2007). This has for instance been shown for experimental participants who have been randomly assigned to represent either plaintiff or defendant (Engel and Glöckner 2013). The effect has been defined as motivated reasoning (Kunda 1990), and for instance been demonstrated for policy preferences regarding discrimination (Braman and Nelson 2007). There is also a cognitive downside. As the mental mechanism is constructive, objective information, like numbers, are liable to inadvertent distortion (Hastie 2011, Rachlinski, Wistrich et al. 2015).

4. Testing Complete Cases

All the former is experimental evidence on individual behavioral effects that one has reason to believe to matter for judicial decision-making. An alternative experimental paradigm exposes real judges to complete materials for a (fictitious) case. That way one for instance finds that judges only pay lip service to precedent when stare decisis holds (Spamann, Klöhn et al. 2021) and take precedent into account when procedural rules forbid (Liu, Klöhn et al. 2021), while being strongly influenced by clearly irrelevant personal characteristics of the defendant (Spamann and Klöhn 2016).

5. Panels

There is a huge literature in political science about decision making within panels of multiple judges (see for instance Songer 1982, Haire, Moyer et al. 2013), some of which is behavioral. It for instance aims at explaining why the presence of a “counterjudge” who has been appointed by the President of a different party (Miles and Sunstein 2006), has different gender (Peresie 2004) or different race (Sommers 2006) tends to lead to more moderate decisions. One explanation is cognitive. The counterjudge alerts the majority of the panel to a concern that they feel unable to overlook (Spitzer and Talley 2013). But relatively little of this evidence is experimental, although group decision making is a classic topic in experimental economics and social psychology (for a structured overview of this literature see Engel 2010).

An exception is Van Dijk, Sonnemans et al. (2014). They show that panels commit less errors than individual judges; that merely aggregating votes suffices; that deliberation does not induce new bias, and improves decision-making if members of the panel decide later cases individually.

6. Fact Finding

The courts do not have the luxury of complete information. Usually facts of the case are unknown, and often additionally contested. Fact finding is an important part of decision-making in the trial courts. Experiments have tested how likely the courts are to get the facts right. A confession may be the result of coercion, rather than the truth. Experimental participants largely ignore the degree of coercion, and instead focus on stereotypes. If the defendant comes from the Middle East and is accused of terrorism, even a dubious confession makes him more likely to be held guilty, same as an African American accused of gang violence, while even a seemingly voluntary confession has little effect if nationality and the charge are matched in the opposite direction (Smalarz, Madon et al. 2018). While judges notice that a confession has been coerced, it still biases the disposition of the case (Wallace and Kassin 2012). If defendants retract a confession, this has practically no effect on its probative value (Malloy and Lamb 2010).

A rich experimental literature has tested the interaction between the court and eyewitnesses. If they express greater confidence, guilty verdicts are substantially more likely (see the meta-study of Slane and Dodson 2022). Students assuming the role of a judge are overly swayed by hearing (mock) eyewitness identification (Boyce, Lindsay et al. 2008). Safeguards routinely implemented by the courts, like cross-examination, prove largely ineffective (see the review of the experimental literature by Devenport, Penrod et al. 1997). Novices and even experts are poor at detecting that a witness lies (Nysse-Carris, Bottoms et al. 2011).

A particularly active area of experimental work concerns lineups. A meta-analysis shows that more innocent fillers do better protect the defendant against false identification (but also increase the number of false negatives) (Juncu and Fitzgerald 2021). If the suspect stands out from the fillers, she is much more likely to be identified, whether innocent or guilty (Fitzgerald, Price et al. 2013). Clear experimental evidence notwithstanding, the courts in the US continue to assign probative value to eyewitness identifications that result from the police suggesting recognition of the suspect (Wells and Quinlivan 2009).

How an expert presents her findings affects how much the judge relies on it. Qualitative statements carry more weight than quantitative statements, as does a statement summarizing the expert's conclusion. Making limitations explicit has no measurable effect (McQuiston-Surrett and Saks 2009). The way in which a mental health expert informs the court about the risk that a defendant found not guilty for insanity is dangerous affects the decision about her release from hospital (Dolores and Redding 2009). The way how DNA evidence is presented dramatically changes how the courts assess its probative value (Koehler 2001). If the expert testifies that she has found supporting evidence, but that it is weak, judicial triers of fact misinterpret this as evidence against the claim (Martire, Kemp et al. 2013). In another experiment, judges were not sensitive to manipulations of the validity and reliability of expert testimony (Chorn and Kovera 2019).

Participants assuming the role of judges who have been biased against a party as part of the experimental design are less biased when exposed to an adversarial, rather than inquisitorial, procedure (Thibaut, Walker et al. 1972). A meta-analysis of experimental studies establishes

a small, but significant reduction in the likelihood of pretrial detention if the trier had access to a formal tool to assess the risk of recidivism (Viljoen, Jonnson et al. 2019). But another experiment found an interaction with race: availability of a risk assessment score helped white, but hurt black defendants (Skeem, Scurich et al. 2020). Experiments demonstrate that judges have a hard time ignoring inadmissible evidence (Lieberman and Arndt 2000, Wistrich, Guthrie et al. 2004). Judges are not much better than jury members at disregarding potentially biasing information that one of the parties has strategically introduced (Landsman and Rakos 1994).

7. Institutional Intervention

Judicial decision-making is institutionally contained. If the legislator expects this to address a normative concern about the carriage of justice, it may redesign elements of the institutional setting. Experiments have tested the impact of multiple of these elements on judicial decision-making, and on case outcomes in particular.

Procedural rules may limit access to information. For instance, lay participants usually do not learn about pretrial information, whereas the professional judges with whom they interact do. This information gap has been shown to put them at a disadvantage in deliberations within the court (Fujita and Hotta 2010).

Procedural law defines which party is heard at which moment of the trial. In criminal procedure, the defendant traditionally has the last word. This gives prosecution the opportunity to set an anchor when justifying the charge (Englich, Mussweiler et al. 2005). However when aggregating the ambiguous information of the case to obtain the final judgement, experiments demonstrate a recency effect: the assessment last heard has the strongest effect on the judgement (Engel, Timme et al. 2020). A recency effect was also found for the order in which multiple pieces of evidence are presented, but only if the decision-maker revalued the evidence after every new piece (Kerstholt and Jackson 1998). A recency effect was also observed if an eyewitness contradicted the prosecution's suspect (Dahl, Brimacombe et al. 2009). Defense attorneys may exploit the recency effect in an attempt at "stealing thunder": they start their pleading with noting the main incriminating evidence, to later dispel it, and leave the court with a more favorable view (Dolnik, Case et al. 2003).

In multiple respects, courts have discretion. The most salient situation is sentencing. One paper interprets the decision of the US Supreme Court in *Booker* as a natural experiment. In that ruling, the Supreme Court had given the trial courts more leeway to depart from the sentence defined by the guidelines (543 U.S. 220 (2005)). The trial courts immediately seized the opportunity, and were more likely to reduce, rather than increase, sentences. Still average sentence length approximately stayed constant, as there were a number of very pronounced upward corrections (Hofer 2007). Bushway, Owens et al. (2012) exploit the fact that, in Maryland, data is available about the true criteria for the state's sentencing guidelines, and the criteria the trial judge has been presented with. In about 10% of the cases, these criteria did not match, i.e. the judge decided on the basis of factually wrong information. Arguing that these errors are random, they find that judges are eager to go along with downward errors, in particular for violent crime. By contrast, judges tend to use discretion to correct upward errors, albeit not fully.

Machine learning and the availability of relatively large datasets make it possible to provide trial judges with machine generated predictions, e.g. about the risk that the defendant will re-convict. Experimental participants are not easily induced to take this machine generated advice into account. In the experiment, they only do if they receive a financial incentive for following the advice. Even a financial incentive for accuracy does not make them rely on the advice (Grgić-Hlača, Engel et al. 2019). In Wisconsin, these computer-generated predictions may also be used for sentencing. The Supreme Court of Wisconsin has cleared this practice (881 [N.W.2d](#) 749 (Wis. 2016)), but has obliged the court administration to add a list of warnings. Experimental subjects largely ignore these warnings. If the warnings are made much more salient and graphic, and if participants are incentivized, warnings have an effect, but it is counterproductive. They follow the machine advice if it has been wrong, while their own prior assessment had been correct (Engel and Grgić-Hlača 2021).

There have been concerns that the presence of still and video cameras in court might have a chilling effect (Osterreicher 2011). In an experiment, these concerns were not supported. In a criminal case, experimental participants were not more likely to find against defendant (Padon 1985).

During the pandemic, witnesses frequently had to be heard over video, rather than in person. An experiment shows that audio quality matters. With poor audio, witnesses are rated as less credible, triers had less precise recollections of their statements, and their testimony had less influence on the case disposition (Bild, Redman et al. 2021).

A natural experiment in Israel made it possible to identify the effect of a reduction in workload on the way how judges handle cases. It turns out that judges not only use the extra leeway to work harder and more diligently on individual cases; in civil cases the extra work is also to the benefit of plaintiffs, who are more likely to win (Engel and Weinshall Margel 2020). In a lab experiment with law students, time pressure increases the probability that participants straightforwardly follow the law. But they are also significantly less confident in the disposition of the case (Sheppard 2011).

For most decisions, judges have to give explicit, often even written reasons. This procedural requirement can potentially have a host of behavioral effects (summarized in Engel 2007). One experiment shows that the requirement may be blunted. Judges may ex post rationalize a decision tainted by bias (Liu and Li 2019). Yet if judges are required to spell out their reasons before deciding, a debiasing effect obtains (Liu 2018).

For the most part, trial judges decide under the shadow of appeal. One behavioral channel is self-esteem (the judge does not want to make a mistake) and social esteem (the judge does not want to lose reputation, by seeing her decision reversed). In an experiment, these effects obtain, and lead to higher accuracy (Feess and Sarel 2018).

8. Formal and Informal Parties

To a considerable degree, judges are in the hands of the parties and their representatives. To understand error and bias in judicial decision-making, one must therefore understand error and bias in what the court hears from the parties. Experiments show that lawyers tend to be overly convinced by the side they represent (Eigen and Listokin 2012). They also make systematic mistakes in the assessment of the evidence, and assign higher subjective probability to events that are described in greater detail (Fox and Birke 2002).

If the police make mistakes, these mistakes are likely to affect the court, directly through hearing the police as witness, or indirectly through prosecution. This is why the experimental research about bias in police investigations is relevant. It has for instance found that police officers tend to be biased against suspects they have themselves arrested (Lidén, Gräns et al. 2018). Racial bias, to the detriment of black defendants, has also been documented (Swencionis and Goff 2017).

Experts tasked with informing the court about the severity of mental health limitations are aware of the risk of bias, but hold the erroneous belief that it can be overcome with willpower alone (Zapf, Kukucka et al. 2018).

9. Public Perception

The judiciary does not have its own enforcement apparatus. This makes the judiciary particularly dependent on public support. Support starts with selection. It is particularly salient for the US Supreme Court. One experiment finds an effect of motivated reasoning (Kunda 1990) on procedural preferences. Those in favor of the candidate wish the Senate hearing to be formal and legalistic. Those opposed to the candidate argue for a more politicized procedure (Badas 2022). A conjoint experiment finds a spillover from the US State to the federal level. Participants from states in which judges are appointed prefer a more legalistic procedure for the selection of Supreme Court justices (Krewson and Owens 2021b). Another conjoint experiment demonstrates that the general public essentially sees US Supreme Court justices as ordinary political actors. Information about their political leanings dominates information about judicial competence (Sen 2017). However, yet another experiment finds that information about judicial philosophies also changes public evaluation of a candidate (Krewson and Owens 2021a).

Learning that judges have received politically motivated support for their campaigns reduces the perceived legitimacy of the court (Gibson, Gottfried et al. 2011). Actually legitimacy is even reduced if the candidate rejects the campaign contribution (Gibson and Caldeira 2012).

Does the general public only care about outcomes, or is trust in courts and their perceived legitimacy also influenced by judicial procedure? Experimental evidence is mixed. One study shows a moderating effect of how the procedure is perceived. If judges are seen as using a principled decision-making process, institutional loyalty determines acceptance. If, by contrast, judges are perceived as policy makers, members of the public only accept the decisions

with which they agree on substance (Woodson 2015). In Norway, where the judiciary is less perceived as politicized, public support increases when the decision comes with a dissenting opinion that makes the normative conflict transparent (Bentsen 2019).

A quasi-experiment from Spain shows that the perception of judicial activism can even reduce support for the entire polity: decisions of the European Court of Justice are perceived as a threat to national sovereignty which, in turn, triggers resistance to European integration (Turnbull-Dugarte and Devine 2022). On the other hand, public support for Veterans Treatment Courts in the US spills over into general support for the legal system. This finding is explained with an effect of procedural justice (Gallagher and Ashford 2021). Support for the judiciary is also increased in participants who are induced to watch a Dutch TV series that makes judicial procedure more transparent (Grimmelikhuijsen and Klijn 2015).

Overall, the effect of the reasons the court gives for a contentious decision had little effect on experimental participants. Priming them with either equality or liberty did not have much effect on them accepting the decision of the US Supreme Court on same-sex marriage. Actually the equality prime even reduced declared acceptance (Cahill and Rapp 2014).

On the opposite end of the spectrum, a court that enjoys high public acceptance, like the German Constitutional Court, even has power to convey legitimacy to political decisions taken by Parliament or Government, or to deprive their decisions of legitimacy (Sternberg, Brouard et al. 2021).

The most important factor for explaining support of experimental participants for a court ruling is the outcome. The judge's mode of reasoning had no measurable effect on support if they agreed with the outcome. If they disagreed with the outcome, acceptance increased if the ruling expressly acknowledged that the opponent also had a point (Simon and Scurich 2011). This result also holds if experimental participants receive information about the case by a legal expert (Simon and Scurich 2013).

III. Pushing the Frontier

Judicial decision-making is definitely not a virgin territory. This survey is rich, and it has only covered experimental research. Should researchers therefore be advised to watch out for alternative topics? Is the expected marginal benefit from new experiments minor? Has the field reached a steady state of normal science? In some dimensions, the answer is probably yes, and this survey may help the next generation of court researchers spotting research questions that they may prefer to avoid. In conclusion I want to sketch areas in which major advances are more likely.

Disciplines have their cultures. Experimental research on judicial decision-making has most intensely engaged with social psychology, and with psychological work on judgement and decision-making. The equally rich literature in behavioral economics has been almost completely neglected, as has the (less rich, but still sizeable) literature on experimental political science.

The extant experimental research on judicial decision-making is surprisingly little legal. Judges decide in a tightly controlled institutional context. There is variance in this context over time, and across jurisdictions. This variance could be productively translated into stylized facts that can make for experimental manipulations.

Most of the experiments reported in this survey aim at identifying a central tendency. They investigate how their sample on average reacts to the manipulation, and infer the treatment effect in the population. This simplification may prove inappropriate. Different groups of judges may react in characteristically different ways: appointed vs. elected; male vs. female; Caucasian vs. of color; with or without attorney experience; with or without expertise in the specific field (like medical malpractice, DNA evidence, or antitrust). It may therefore be a promising next step to design experiments that are able to document patterned heterogeneity, and the resulting heterogeneous treatment effects.

Finally no case is like the next, and as a practical matter, court cases are only incompletely defined. Traditional experimental methods deliberately eschew this challenge, and purposefully simplify the setting. As machine learning methods are powerful ways of organizing such rich datasets, next generation experimental research might introduce controlled manipulations into such more naturalistic settings, and might rely on artificial intelligence to organize the resulting data. Such experiments would have only partly ex ante defined dependent variables, and would blend theory driven interventions with pattern finding.

Experimental research on judicial decision-making is likely to thrive!

References

- ANDERSON, JOHN C, MARIANNE M JENNINGS, D JORDAN LOWE AND PHILIP MJ RECKERS (1997). "The Mitigation of Hindsight Bias in Judges' Evaluation of Auditor Decisions." Auditing **16**(2): 20-39.
- ARIAS, ERIC (2019). Impartiality in international courts: Evidence from a natural experiment at the WTO. https://www.peio.me/wp-content/uploads/2019/01/PEIO12_paper_102.pdf.
- BADAS, ALEX (2022). "Motivated Reasoning and Attitudes Towards Supreme Court Confirmation Hearings. Evidence from Five Nominations and an Experiment." Political Research Quarterly *******: *******.
- BENTSEN, HENRIK LITLERÉ (2019). "Dissent, Legitimacy, and Public Support for Court Decisions. Evidence from a Survey-Based Experiment." Law and Society Review **53**(2): 588-610.
- BILD, ELENA, ANNABEL REDMAN, ERYN J NEWMAN, BETHANY R MUIR, DAVID TAIT AND NORBERT SCHWARZ (2021). "Sound and Credibility in the Virtual Court. Low Audio Quality Leads to Less Favorable Evaluations of Witnesses and Lower Weighting of Evidence." Law and Human Behavior **45**(5): 481-495.
- BORDALO, PEDRO, NICOLA GENNAIOLI AND ANDREI SHLEIFER (2015). "Salience Theory of Judicial Decisions." Journal of Legal Studies **44**(S1): S7-S33.
- BORNSTEIN, BRIAN H, JONATHAN M GOLDING, JEFFREY NEUSCHATZ, CHRISTOPHER KIMBROUGH, KRYSIA REED, CASEY MAGYARICS AND KATHERINE LUECHT (2017). "Mock Juror Sampling Issues in Jury Simulation Research. A Meta-Analysis." Law and Human Behavior **41**(1): 13-28.
- BOYCE, MELISSA A, D STEPHEN LINDSAY AND CAE BRIMACOMBE (2008). "Investigating Investigators. Examining the Impact of Eyewitness Identification Evidence on Student-Investigators." Law and Human Behavior **32**(5): 439-453.
- BRAMAN, EILEEN (2006). "Reasoning on the Threshold. Testing the Separability of Preferences in Legal Decision Making." Journal of Politics **68**(2): 308-321.
- BRAMAN, EILEEN AND THOMAS E NELSON (2007). "Mechanism of Motivated Reasoning? Analogical Perception in Discrimination Disputes." American Journal of Political Science **51**(4): 940-956.
- BUSHWAY, SHAWN D, EMILY G OWENS AND ANNE MORRISON PIEHL (2012). "Sentencing Guidelines and Judicial Discretion. Quasi-Experimental Evidence From Human Calculation Errors." Journal of Empirical Legal Studies **9**(2): 291-319.
- BYSTRANOWSKI, PIOTR, BARTOSZ JANIK, MACIEJ PRÓCHNICKI AND PAULINA SKÓRSKA (2021). "Anchoring Effect in Legal Decision-Making. A Meta-Analysis." Law and Human Behavior **45**(1): 1-23.

- CAHILL, COURTNEY MEGAN AND GEOFFREY CHRISTOPHER RAPP (2014). "Does the Public Care How the Supreme Court Reasons. Empirical Evidence from a National Experiment and Normative Concerns in the Case of Same-Sex Marriage." North Carolina Law Review **93**: 303-374.
- CERRONE, CLAUDIA AND CHRISTOPH ENGEL (2019). "Deciding on Behalf of Others Does Not Mitigate Selfishness: An Experiment." Economics Letters **183**: 108616.
- CHEN, DANIEL L, TOBIAS J MOSKOWITZ AND KELLY SHUE (2016). "Decision Making Under the Gambler's Fallacy. Evidence From Asylum Judges, Loan Officers, and Baseball Umpires." Quarterly Journal of Economics **131**(3): 1181-1242.
- CHOI, DONGHYUN DANNY, J. ANDREW HARRIS AND FIONA SHEN-BAYH (2022). "Ethnic Bias in Judicial Decision Making: Evidence from Criminal Appeals in Kenya." American Political Science Review **116**: 1067-1080.
- CHORN, JACQUELINE AUSTIN AND MARGARET BULL KOVERA (2019). "Variations in Reliability and Validity Do Not Influence Judge, Attorney, and Mock Juror Decisions About Psychological Expert Evidence." Law and Human Behavior **43**(6): 542-557.
- DAHL, LEORA C, CA BRIMACOMBE AND D STEPHEN LINDSAY (2009). "Investigating Investigators. How Presentation Order Influences Participant–Investigators' Interpretations of Eyewitness Identification and Alibi Evidence." Law and Human Behavior **33**(5): 368-380.
- DESMARAIS, SARAH L AND J DON READ (2011). "After 30 Years, What Do We Know About What Jurors Know? A Meta-Analytic Review of Lay Knowledge Regarding Eyewitness Factors." Law and Human Behavior **35**(3): 200-210.
- DEVENPORT, JENNIFER L, STEVEN D PENROD AND BRIAN L CUTLER (1997). "Eyewitness Identification Evidence. Evaluation Commonsense Evaluations." Psychology, Public Policy, and Law **3**(2-3): 338-361.
- DEVINE, DENNIS J, LAURA D CLAYTON, BENJAMIN B DUNFORD, RASMY SEYING AND JENNIFER PRYCE (2001). "Jury Decision Making. 45 Years of Empirical Research on Deliberating Groups." Psychology, Public Policy, and Law **7**(3): 622-727.
- DOLNIK, LARA, TREVOR I CASE AND KIPLING D WILLIAMS (2003). "Stealing Thunder as a Courtroom Tactic Revisited. Processes and Boundaries." Law and Human Behavior **27**(3): 267-287.
- DOLORES, JOHN C AND RICHARD E REDDING (2009). "The Effects of Different Forms of Risk Communication on Judicial Decision Making." International Journal of Forensic Mental Health **8**(2): 142-146.
- EIGEN, ZEV J AND YAIR LISTOKIN (2012). "Do Lawyers Really Believe Their Own Hype, and Should They? A Natural Experiment." Journal of Legal Studies **41**(2): 239-267.

- EISENBERG, THEODORE, JEFFREY J RACHLINSKI AND MARTIN T WELLS (2001). "Reconciling Experimental Incoherence With Real-World Coherence in Punitive Damages." Stanford Law Review **54**: 1239-1271.
- ENGEL, CHRISTOPH (2007). The Psychological Case for Obliging Judges to Write Reasons. The Impact of Court Procedure on the Psychology of Judicial Decision Making. C. Engel and F. Strack. Baden-Baden, Nomos: 71-109.
- ENGEL, CHRISTOPH (2010). "The Behaviour of Corporate Actors. A Survey of the Empirical Literature." Journal of Institutional Economics **6**: 445-475.
- ENGEL, CHRISTOPH AND ANDREAS GLÖCKNER (2013). "Role Induced Bias in Court. An Experiment." Journal of Behavioral Decision Making **26**: 272-284.
- ENGEL, CHRISTOPH AND SEBASTIAN J GOERG (2018). "If the Worst Comes to the Worst. Dictator Giving when Recipient's Endowments are Risky." European Economic Review **105**: 51-70.
- ENGEL, CHRISTOPH AND NINA GRGIC-HLACA (2021). "Machine Advice with a Warning about Machine Limitations. Experimentally Testing the Solution Mandated by the Wisconsin Supreme Court." Journal of Legal Analysis **13**: 284-340.
- ENGEL, CHRISTOPH, SINIKA TIMME AND ANDREAS GLÖCKNER (2020). "Coherence-Based Reasoning and Order Effects in Legal Judgments." Psychology, Public Policy, and Law *******: *******.
- ENGEL, CHRISTOPH AND KEREN WEINSHALL MARGEL (2020). "Manna from Heaven for Judges. Judges' Reaction to a Quasi-Random Reduction in Caseload." Journal of Empirical Legal Studies **17**: 722-751.
- ENGEL, CHRISTOPH AND LILIA ZHURAKHOVSKA (2017). "You Are In Charge. Experimentally Testing the Motivating Power of Holding a Judicial Office." Journal of Legal Studies **46**: 1-50.
- ENGLISH, BIRTE, THOMAS MUSSWEILER AND FRITZ STRACK (2005). "The Last Word in Court. A Hidden Disadvantage for the Defense." Law and Human Behavior **29**(6): 705-722.
- EPSTEIN, LEE, JACK KNIGHT AND ANDREW D MARTIN (2001). "The Supreme Court as a Strategic National Policymaker." Emory Law Journal **50**: 583-612.
- FEESS, EBERHARD AND ROEE SAREL (2018). "Judicial Effort and the Appeals System. Theory and Experiment." Journal of Legal Studies **47**(2): 269-294.
- FEIGENSON, NEAL AND JAIHYUN PARK (2006). "Emotions and Attributions of Legal Responsibility and Blame. A Research Review." Law and Human Behavior **30**(2): 143-161.
- FITZGERALD, RYAN J, HEATHER L PRICE, CHRIS ORIET AND STEVE D CHARMAN (2013). "The Effect of Suspect-Filler Similarity on Eyewitness Identification Decisions. A Meta-Analysis." Psychology, Public Policy, and Law **19**(2): 151-164.

- FOX, CRAIG R AND RICHARD BIRKE (2002). "Forecasting Trial Outcomes. Lawyers Assign Higher Probability to Possibilities That Are Described in Greater Detail." Law and Human Behavior **26**(2): 159-173.
- FUJITA, MASAHIRO AND SYUGO HOTTA (2010). "The Impact of Differential Information Between Lay Participants and Professional Judges on Deliberative Decision-Making." International Journal of Law, Crime and Justice **38**(4): 216-235.
- FUNK, FRIEDERIKE AND ALEXANDER TODOROV (2013). "Criminal Stereotypes in the Courtroom. Facial Tattoos Affect Guilt and Punishment Differently." Psychology, Public Policy, and Law **19**(4): 466-478.
- FURGESON, JOSHUA R, LINDA BABCOCK AND PETER M SHANE (2008). "Do a Law's Policy Implications Affect Beliefs About Its Constitutionality? An Experimental Test." Law and Human Behavior **32**(3): 219-227.
- GALLAGHER, JOHN M AND JOSÉ B ASHFORD (2021). "Perceptions of Legal Legitimacy in Veterans Treatment Courts. A Test of a Modified Version of Procedural Justice Theory." Law and Human Behavior **45**(2): 152-164.
- GAZAL-AYAL, OREN AND RAANAN SULITZEANU-KENAN (2010). "Let My People Go. Ethnic in-Group Bias in Judicial Decisions—Evidence From a Randomized Natural Experiment." Journal of Empirical Legal Studies **7**(3): 403-428.
- GERBASI, KATHLEEN C, MIRON ZUCKERMAN AND HARRY T REIS (1977). "Justice Needs a New Blindfold. A Review of Mock Jury Research." Psychological Bulletin **84**(2): 323-345.
- GIBSON, JAMES L AND GREGORY A CALDEIRA (2012). "Campaign Support, Conflicts of Interest, and Judicial Impartiality. Can Recusals Rescue the Legitimacy of Courts?" Journal of Politics **74**(1): 18-34.
- GIBSON, JAMES L, JEFFREY A GOTTFRIED, MICHAEL X DELLI CARPINI AND KATHLEEN HALL JAMIESON (2011). "The Effects of Judicial Campaign Activity on the Legitimacy of Courts. A Survey-Based Experiment." Political Research Quarterly **64**(3): 545-558.
- GIRVAN, ERIK J (2016). "Wise Restraints?: Learning Legal Rules, Not Standards, Reduces the Effects of Stereotypes in Legal Decision-Making." Psychology, Public Policy, and Law **22**(1): 31-45.
- GLÖCKNER, ANDREAS AND CHRISTOPH ENGEL (2013). "Can We Trust Intuitive Jurors? Standards of Proof and the Probative Value of Evidence in Coherence Based Reasoning." Journal of Empirical Legal Studies **10**: 230-252.
- GRGIĆ-HLAČA, NINA, CHRISTOPH ENGEL AND KRISHNA P GUMMADI (2019). Human Decision Making with Machine Assistance. An experiment on Bailing and Jailing. Proceedings of the ACM on Human-Computer Interaction.

- GRIMMELIKHUIJSEN, STEPHAN AND ALBERT KLIJN (2015). "The Effects of Judicial Transparency on Public Trust. Evidence From a Field Experiment." Public Administration **93**(4): 995-1011.
- GUTHRIE, CHRIS, JEFFREY J. RACHLINSKI AND ANDREW J. WISTRICH (2001). "Inside the Judicial Mind." Cornell Law Review **86**: 777-830.
- GUTHRIE, CHRIS, JEFFREY J. RACHLINSKI AND ANDREW J. WISTRICH (2007). "Blinking on the Bench. How Judges Decide Cases." Cornell Law Review **93**: 1-43.
- HAIRE, SUSAN B, LAURA P MOYER AND SHAWN TREIER (2013). "Diversity, Deliberation, and Judicial Opinion Writing." Journal of Law and Courts **1**(2): 303-330.
- HASTIE, REID (2011). "The Challenge to Produce Useful Legal Numbers." Journal of Empirical Legal Studies **8**: 6-20.
- HASTIE, REID, STEVEN PENROD AND NANCY PENNINGTON (1983). Inside the Jury. Cambridge, Mass., Harvard University Press.
- HOETGER, LORI A, DENNIS J DEVINE, EVE M BRANK, RYAN M DREW AND REBECCA REES (2022). "The Impact of Pretrial Publicity on Mock Juror and Jury Verdicts. A Meta-Analysis." Law and Human Behavior **46**: 121-139.
- HOFER, PAUL J (2007). "United States v. Booker as a Natural Experiment. Using Empirical Research to Inform the Federal Sentencing Policy Debate." Criminology and Public Policy **6**(3): 433-460.
- HOLYOAK, KEITH J. AND DAN SIMON (1999). "Bidirectional Reasoning in Decision Making by Constraint Satisfaction." Journal of Experimental Psychology: General **128**: 1-29.
- JUNCU, STEFANA AND RYAN J FITZGERALD (2021). "A Meta-Analysis of Lineup Size Effects on Eyewitness Identification." Psychology, Public Policy, and Law **27**: 295-315.
- KAHAN, DAN M, DAVID HOFFMAN, DANIELI EVANS, NEAL DEVINS, EUGENE LUCCI AND KATHERINE CHENG (2015). "Ideology or Situation Sense. An Experimental Investigation of Motivated Reasoning and Professional Judgment." University of Pennsylvania Law Review **164**: 349-439.
- KANTOROWICZ-REZNICHENKO, ELENA, JAROSŁAW KANTOROWICZ AND KEREN WEINSHALL (2022). "Ideological Bias in Constitutional Judgments. Experimental Analysis and Potential Solutions." Journal of Empirical Legal Studies *******: *******.
- KELMAN, MARK, YUVAL ROTTENSTREICH AND AMOS TVERSKY (1996). "Context-Dependence in Legal Decision Making." Journal of Legal Studies **25**(2): 287-318.
- KERSTHOLT, JOSÉ H AND JANET L JACKSON (1998). "Judicial Decision Making. Order of Evidence Presentation and Availability of Background Information." Applied Cognitive Psychology **12**(5): 445-454.

- KOEHLER, JONATHAN J (2001). "When Are People Persuaded By DNA Match Statistics?" Law and Human Behavior **25**(5): 493-513.
- KREWSON, CHRISTOPHER N AND RYAN J OWENS (2021a). "Public Support for Judicial Philosophies. Evidence from a Conjoint Experiment." Journal of Law and Courts **9**(1): 89-110.
- KREWSON, CHRISTOPHER AND RYAN J OWENS (2021b). "How State Judicial Selection Methods May Influence Views of US Supreme Court Nominees. Evidence from a Conjoint Experiment." Journal of Law and Courts ***: ***.
- KULIK, CAROL T, ELISSA L PERRY AND MOLLY B PEPPER (2003). "Here Comes the Judge. The Influence of Judge Personal Characteristics on Federal Sexual Harassment Case Outcomes." Law and Human Behavior **27**(1): 69-86.
- KUNDA, ZIVA (1990). "The Case for Motivated Reasoning." Psychological Bulletin **108**: 480-498.
- KUNDA, ZIVA (1999). Social Cognition. Making Sense of People. Cambridge, Mass., MIT Press.
- LANDSMAN, STEPHAN AND RICHARD F RAKOS (1994). "A Preliminary Inquiry Into the Effect of Potentially Biasing Information on Judges and Jurors in Civil Litigation." Behavioral Sciences & the Law **12**(2): 113-126.
- LIDÉN, MOA, MINNA GRÄNS AND PETER JUSLIN (2018). "The Presumption of Guilt in Suspect Interrogations. Apprehension as a Trigger of Confirmation Bias and Debiasing Techniques." Law and Human Behavior **42**(4): 336-354.
- LIEBERMAN, JOEL D AND JAMIE ARNDT (2000). "Understanding the Limits of Limiting Instructions. Social Psychological Explanations for the Failures of Instructions to Disregard Pretrial Publicity and Other Inadmissible Evidence." Psychology, Public Policy, and Law **6**(3): 677-711.
- LIU, JOHN ZHUANG, LARS KLÖHN AND HOLGER SPAMANN (2021). "Precedents and Chinese Judges. An Experiment." American Journal of Comparative Law **69**(1): 93-135.
- LIU, JOHN ZHUANG AND XUEYAO LI (2019). "Legal Techniques for Rationalizing Biased Judicial Decisions. Evidence From Experiments With Real Judges." Journal of Empirical Legal Studies **16**(3): 630-670.
- LIU, ZHUANG (2018). "Does Reason Writing Reduce Decision Bias? Experimental Evidence From Judges in China." Journal of Legal Studies **47**(1): 83-118.
- LUCAS, JEFFREY W (2003). "Theory-testing, Generalization, and the Problem of External Validity." Sociological Theory **21**(3): 236-253.
- MACCOUN, ROBERT J (1989). "Experimental Research on Jury Decision-Making." Science **244**(4908): 1046-1050.

- MACCOUN, ROBERT J (1990). "Experimental Research on Jury Decision-Making." Jurimetrics **30**: 223-233.
- MALLOY, LINDSAY C AND MICHAEL E LAMB (2010). "Biases in Judging Victims and Suspects Whose Statements Are Inconsistent." Law and Human Behavior **34**(1): 46-48.
- MARTIRE, KRISTY A, RICHARD I KEMP, IAN WATKINS, MALINDI A SAYLE AND BEN R NEWELL (2013). "The Expression and Interpretation of Uncertain Forensic Science Evidence. Verbal Equivalence, Evidence Strength, and the Weak Evidence Effect." Law and Human Behavior **37**(3): 197-207.
- McFATTER, ROBERT M (1978). "Sentencing Strategies and Justice. Effects of Punishment Philosophy on Sentencing Decisions." Journal of Personality and Social Psychology **36**(12): 1490-1500.
- MCQUISTON-SURRETT, DAWN AND MICHAEL J SAKS (2009). "The Testimony of Forensic Identification Science. What Expert Witnesses Say and What Factfinders Hear." Law and Human Behavior **33**(5): 436-453.
- MILES, THOMAS J AND CASS R SUNSTEIN (2006). "Do Judges Make Regulatory Policy? An Empirical Investigation of Chevron." University of Chicago Law Review: 823-881.
- MILLER, ANDREA L (2019). "Expertise Fails to Attenuate Gendered Biases in Judicial Decision-Making." Social Psychological and Personality Science **10**(2): 227-234.
- MITCHELL, TARA L, RYANN M HAW, JEFFREY E PFEIFER AND CHRISTIAN A MEISSNER (2005). "Racial Bias in Mock Juror Decision-Making. A Meta-Analytic Review of Defendant Treatment." Law and Human Behavior **29**(6): 621-637.
- NORTON, MICHAEL I, SAMUEL R SOMMERS, JOSEPH A VANDELLO AND JOHN M DARLEY (2006). "Mixed Motives and Racial Bias. The Impact of Legitimate and Illegitimate Criteria on Decision Making." Psychology, Public Policy, and Law **12**(1): 36-55.
- NYSSE-CARRIS, KARI L, BETTE L BOTTOMS AND JESSICA M SALERNO (2011). "Experts' and Novices' Abilities to Detect Children's High-Stakes Lies of Omission." Psychology, Public Policy, and Law **17**(1): 76-98.
- OEBERST, AILEEN AND INGKE GOECKENJAN (2016). "When Being Wise After the Event Results in Injustice. Evidence for Hindsight Bias in Judges' Negligence Assessments." Psychology, Public Policy, and Law **22**(3): 271-279.
- OSTERREICHER, MICKEY H (2011). "Cameras in the Courts. The Long Road to the New Federal Experiment." Reynolds Court and Media Law Journal **1**: 221-258.
- PADDON, ANNA R (1985). Television Coverage of Criminal Trials With Cameras and Microphones. A Laboratory Experiment of Audience Effects (Videotape, Photograph, Court, Fair Trial). <https://www.proquest.com/pagepdf/303422540?accountid=16589>.

- PENNINGTON, NANCY AND REED HASTIE (1986). "Evidence Evaluation in Complex Decision Making." Journal of Personality and Social Psychology **51**: 242-258.
- PENNINGTON, NANCY AND REED HASTIE (1991). "A Cognitive Theory of Juror Decision Making. The Story Model." Cardozo Law Review **13**: 519-557.
- PERESIE, JENNIFER L (2004). "Female Judges Matter. Gender and Collegial Decisionmaking in the Federal Appellate Courts." Yale Law Journal **114**: 1759-1790.
- POSNER, RICHARD A. (1993). "What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)." Supreme Court Economic Review **3**: 1-41.
- RACHLINSKI, JEFFREY J, CHRIS GUTHRIE AND ANDREW J WISTRICH (2006). "Inside the Bankruptcy Judge's Mind." Boston University Law Review **86**: 1227-1265.
- RACHLINSKI, JEFFREY J, SHERI LYNN JOHNSON, ANDREW J WISTRICH AND CHRIS GUTHRIE (2008). "Does Unconscious Racial Bias Affect Trial Judges." Notre Dame Law Review **84**: 1195-1246.
- RACHLINSKI, JEFFREY J, ANDREW J WISTRICH AND CHRIS GUTHRIE (2015). "Can Judges Make Reliable Numeric Judgments. Distorted Damages and Skewed Sentences." Indiana Law Journal **90**: 695-739.
- ROSE, MARY R, JANICE NADLER AND JIM CLARK (2006). "Appropriately Upset? Emotion Norms and Perceptions of Crime Victims." Law and Human Behavior **30**(2): 203-219.
- SCHRAM, ARTHUR (2005). "Artificiality. The Tension Between Internal and External Validity in Economic Experiments." Journal of Economic Methodology **12**(2): 225-237.
- SCURICH, NICHOLAS AND RICHARD S JOHN (2011). "The Effect of Framing Actuarial Risk Probabilities on Involuntary Civil Commitment Decisions." Law and Human Behavior **35**(2): 83-91.
- SEGAL, JEFFREY ALLAN AND HAROLD J. SPAETH (1993). The Supreme Court and the Attitudinal Model. Cambridge ; New York, NY, Cambridge University Press.
- SEN, MAYA (2017). "How Political Signals Affect Public Support for Judicial Nominations. Evidence from a Conjoint Experiment." Political Research Quarterly **70**(2): 374-393.
- SHEPHERD, JOANNA M. (2009). "The Influence of Retention Politics on Judges' Voting." Journal of Legal Studies **38**: 169-206.
- SHEPPARD, BRIAN (2011). "Judging Under Pressure. A Behavioral Examination of the Relationship Between Legal Decisionmaking and Time." Florida State University Law Review **39**: 931-1001.
- SIMON, DAN (1998). "A Psychological Model of Judicial Decision Making." Rutgers Law Journal **30**: 1-142.

- SIMON, DAN (2004). "A Third View of the Black Box. Cognitive Coherence in Legal Decision Making." University of Chicago Law Review **71**: 511-586.
- SIMON, DAN (2019). "On Juror Decision Making. An Empathic Inquiry." Annual Review of Law and Social Science **15**: 415-435.
- SIMON, DAN, LIEN B. PHAM, QUANG A. LE AND KEITH J. HOLYOAK (2001). "The Emergence of Coherence Over the Course of Decision Making." Journal of Experimental Psychology: Learning, Memory and Cognition **27**: 1250-1260.
- SIMON, DAN AND NICHOLAS SCURICH (2011). "Lay judgments of judicial decision making." Journal of Empirical Legal Studies **8**(4): 709-727.
- SIMON, DAN AND NICHOLAS SCURICH (2013). "The Effect of Legal Expert Commentary on Lay Judgments of Judicial Decision Making." Journal of Empirical Legal Studies **10**(4): 797-814.
- SKEEM, JENNIFER, NICHOLAS SCURICH AND JOHN MONAHAN (2020). "Impact of Risk Assessment on Judges' Fairness in Sentencing Relatively Poor Defendants." Law and Human Behavior **44**(1): 51-59.
- SLANE, CRYSTAL R AND CHAD S DODSON (2022). "Eyewitness Confidence and Mock Juror Decisions of Guilt. A Meta-Analytic Review." Law and Human Behavior **46-66**(1): 45.
- SMALARZ, LAURA, STEPHANIE MADON AND ANNA TUROSAK (2018). "Defendant Stereotypicality Moderates the Effect of Confession Evidence on Judgments of Guilt." Law and Human Behavior **42**(4): 355-368.
- SOMMERS, SAMUEL R (2006). "On Racial Diversity and Group Decision Making. Identifying Multiple Effects of Racial Composition on Jury Deliberations." Journal of Personality and Social Psychology **90**(4): 597-612.
- SONGER, DONALD R (1982). "Consensual and Nonconsensual Decisions in Unanimous Opinions of the United States Courts of Appeals." American Journal of Political Science **26**: 225-239.
- SONNEMANS, JOEP AND FRANS VAN DIJK (2012). "Errors in Judicial Decisions: Experimental Results." Journal of Law, Economics, and Organization **28**(4): 687-716.
- SPAMANN, HOLGER AND LARS KLÖHN (2016). "Justice is Less Blind, and Less Legalistic, Than We Thought: Evidence from an Experiment with Real Judges." Journal of Legal Studies **45**(2): 255-280.
- SPAMANN, HOLGER, LARS KLÖHN, CHRISTOPHE JAMIN, VIKRAMADITYA KHANNA, JOHN ZHUANG LIU, PAVAN MAMIDI, ALEXANDER MORELL AND IVAN REIDEL (2021). "Judges in the Lab: No Precedent Effects, No Common/Civil Law Differences." Journal of Legal Analysis **13**(1): 110-126.

- SPITZER, MATTHEW AND ERIC TALLEY (2013). "Left, Right, and Center. Strategic Information Acquisition and Diversity in Judicial Panels." Journal of Law, Economics, & Organization **29**(3): 638-680.
- STEBLAY, NANCY, HARMON M. HOSCH, SCOTT E. CULHANE AND ADAM MCWETHY (2006). "The Impact on Juror Verdicts of Judicial Instruction to Disregard Inadmissible Evidence: A Meta-Analysis." Law and Human Behavior **30**: 469-492.
- STEBLAY, NANCY MEHRKENS, JASMINA BESIREVIC, SOLOMON M FULERO AND BELIA JIMENEZ-LORENTE (1999). "The Effects of Pretrial Publicity on Juror Verdicts. A Meta-analytic Review." Law and Human Behavior **23**(2): 219-235.
- STERNBERG, SEBASTIAN, SYLVAIN BROUARD AND CHRISTOPH HÖNNIGE (2021). "The legitimacy-conferring capacity of constitutional courts: Evidence from a comparative survey experiment." European Journal of Political Research.
- SWENCIONIS, JILLIAN K AND PHILLIP ATIBA GOFF (2017). "The Psychological Science of Racial Bias and Policing." Psychology, Public Policy, and Law **23**(4): 398-409.
- TEICHMAN, DORON AND EYAL ZAMIR (2014). Judicial Decision-Making. A Behavioral Perspective. Oxford Handbook of Behavioral Economics and the Law. E. Zamir and D. Teichman. Oxford, Oxford University Press: 664-702.
- THIBAUT, JOHN, LAURENS WALKER AND E ALLAN LIND (1972). "Adversary Presentation and Bias in Legal Decisionmaking." Harvard Law Review **86**: 386-401.
- TURNBULL-DUGARTE, STUART J AND DANIEL DEVINE (2022). "Can EU Judicial Intervention Increase Polity Scepticism? Quasi-Experimental Evidence From Spain." Journal of European Public Policy **29**(6): 865-890.
- TYLER, TOM R. (2006). Why People Obey the Law. New Haven, Yale University Press.
- VAN DIJK, FRANS, JOEP SONNEMANS AND EDDY BAUW (2014). "Judicial Error by Groups and Individuals." Journal of Economic Behavior & Organization **108**: 224-235.
- VILJOEN, JODI L, MELISSA R JOHNSON, DANA M COCHRANE, LEE M VARGEN AND GINA M VINCENT (2019). "Impact of Risk Assessment Instruments on Rates of Pretrial Detention, Postconviction Placements, and Release. A Systematic Review and Meta-Analysis." Law and Human Behavior **43**(5): 397-420.
- WALLACE, D BRIAN AND SAUL M KASSIN (2012). "Harmless Error Analysis: How Do Judges Respond to Confession Errors?" Law and Human Behavior **36**(2): 151-157.
- WELLS, GARY L AND DEAH S QUINLIVAN (2009). "Suggestive Eyewitness Identification Procedures and the Supreme Court's Reliability Test in Light of Eyewitness Science. 30 Years Later." Law and Human Behavior **33**(1): 1-24.

- WIENER, RICHARD L, DANIEL A KRAUSS AND JOEL D LIEBERMAN (2011). "Mock Jury Research. Where Do We Go From Here?" Behavioral Sciences and the Law **29**(3): 467-479.
- WISTRICH, ANDREW J, CHRIS GUTHRIE AND JEFFREY J RACHLINSKI (2004). "Can Judges Ignore Inadmissible Information. The Difficulty of Deliberately Disregarding." University of Pennsylvania Law Review **153**: 1251-1346.
- WISTRICH, ANDREW J, JEFFREY J RACHLINSKI AND CHRIS GUTHRIE (2014). "Heart Versus Head. Do Judges Follow the Law of Follow Their Feelings." Texas Law Review **93**: 855-923.
- WOODSON, BENJAMIN (2015). "Politicization and the Two Modes of Evaluating Judicial Decisions." Journal of Law and Courts **3**(2): 193-221.
- ZAPF, PATRICIA A, JEFF KUKUCKA, SAUL M KASSIN AND ITIEL E DROR (2018). "Cognitive Bias in Forensic Mental Health Assessment. Evaluator Beliefs About Its Nature and Scope." Psychology, Public Policy, and Law **24**(1): 1-10.