Editorial Board

Professor Tania Sourdin
Dean and Head of School, Newcastle Law School

Professor Christoph Antons
Professor, Newcastle Law School

Professor Lisa Toohey
LLB Program Convenor, Newcastle Law School

Associate Professor John Anderson
JD Program Convenor, Newcastle Law School

Mr Shaun McCarthy
Director, The University of Newcastle Legal Centre
# Table of Contents

**Editorial**

– *Tania Sourdin and Bin Li*

**Special Contribution**

Language: The Law’s Essential Tool  
– *The Hon Justice Margaret J. Beazley AO*

**Articles**

Evidence-based law and practice with disadvantaged populations: Perspectives of parents involved in child protection proceedings  
– *Nicola Ross, Lou Johnston, Jessica Cocks and Lynette Stoker*

Methods to Evaluate Justice Practices in Eliciting Evidence from Complainants of Child Sexual Abuse  
– *Jane Goodman-Delahunty, Eunro Lee, Martine B Powell and Nina Westera*

Empirical Methods in Law and Psychology: A Review Note on the Research Fostering the New Korean Jury System  
– *Eunro Lee, Yuhwa Han, Kwangbai Park, Sangjoon Kim and Minchi Kim*

Using Pretest-Posttest Research Designs to Enhance Jury Decisions  
– *Jane Goodman-Delahunty and Natalie Martschuk*

Open Data: Turning Data into Information, and Information into Insights that Allow for Evidence-Based Policy  
– *Alana Maurushat, Jan Traskowski and Hadeel Al-Alosi*

**Research Note**

Effectiveness in Mediation: a New Approach  
– *Alysoun Boyle*
Editorial

The Newcastle Law Review, the journal of Newcastle Law School, was first launched in 1995. Since that time an impressive series of articles and notes has been published, representing traditional categories of legal scholarship as well as interdisciplinary contributions. As the leading clinical law school in Australia, that focuses on clinical and interdisciplinary approaches to teaching, the research focus within Newcastle Law School is diverse although grouped within four main theme areas: international law; law and society; applied law and justice, and justice innovation.

As a very active research School, much of our research is underpinned by deep engagement within our local community and beyond. Newcastle Law School also aims to have an impact. In this regard, Newcastle Law School plays a pivotal role in promoting access to justice. This is achieved through engagement with local, state, national, and international communities; academic, as well as applied and evidence-based research focused on improving the capability and capacity of the justice system and minimising barriers to access; the promotion of clarity and fairness through public interest advocacy and test cases; and other clinical work performed by the University of Newcastle Legal Centre.

The reinvigoration of the Newcastle Law Review with a broader circulation is part of our commitment to contribute to dialogue and discussion by supporting from 2018 an open access approach. From this edition, the Newcastle Law Review will be available free of charge to a global readership with the format of online open-access resource. Our publication supports double blind peer reviewing and also has an editorial Board to oversee the operations and content of each issue.

It is also intended that Newcastle Law Review will reflect some of the significant outreach activities undertaken by Newcastle Law School. Whilst traditional academic events that include the Hon Michael Kirby Lecture and the Sir Ninian Stephen Lecture (now in its 26th year) continue, the Newcastle Law School also hosts a number of symposia, conferences, academic and professional seminars and higher degree by research focused events. One commitment of Newcastle to research in 2017 was to host a ‘Symposium on Evidence-based Law and Practice’. The feasibility of such an event was secured by external and internal grants that had been approved to support Newcastle researchers to investigate in the area of evidence-based law. However, the underlying reason for Newcastle to commit to research in this particular field is that we have seen the value of evidence-based approach to legal scholarship and evidence-based approaches to law and justice research underpins much of the research undertaken at Newcastle Law School.¹ This issue of The Newcastle Law Review focusses on evidence-based law and practice and shares the fruits of the Symposium with our readers.

The concept of evidence-based thinking arguably originates from the medical area where the term of ‘evidence-based medicine’ emerged and was developed. As David Sackett and others wrote in their article that ‘The practice of evidence based medicine means integrating

individual clinical expertise with the best available external clinical evidence from systematic research.' 2 The authors argued that even the most experienced clinicians needed external systematic research evidence to support them in making best decisions as to patient care. In other words, good medical service can only be achieved based on systematic research and scientific evidence, rather than on the experience of the clinicians alone.

More recently, the evidence-based approach has been transplanted by business academics to describe the movement in business area where people subject even their deeply-held assumptions to empirical evidence. In his paper, Professor Rachlinski raised an interesting example where in the USA, Wal-Mart stores sold turkey in July as sales data had shown that customers had the same desire for turkey meat in both summer and winter. This example sheds light on how these business operators began to question their belief that turkey would only be popular whilst customers in winter. 3 The Wal-Mart’s decision to sell turkey in July is based on the prior collection of and analysis on its sales data, or ‘evidence’ from its empirical analysis of this data.

Apart from the application of evidence-based approach in business, the evidence-based movement has also been increasingly relevant in other areas of social sciences, including the political science area. Specifically, from the perspective of policy-making, evidence-based methodology has been endorsed and embraced by many countries, including Australia. In 2008, the then Australian Prime Minister Kevin Rudd noted, in his Address to Heads of Agencies and Members of Senior Executive Service 4, that one of the elements of Australian Government's vision for the future Australian public service would be ‘developing evidence-based policy making processes as part of a robust culture of policy contestability’. He further emphasized the importance of the Government to ‘receiv[ing] the best advice, based on the best available information and evidence.’

The notion of evidence-based policy clearly contributed to the formation of 2013 Report on ‘Science for Policy: Mapping Australian Government Investments and Institutions Discussion’ 5 where the effect of the evidence-based policy on public service reform was examined and explored. The efforts of the Australian Government, academia and business have in recent years further cemented the notion that an evidence base is critical in terms of the policy making process in Australia and how Government decisions about the public sector can be made. Interestingly, recent criticism on some of government policies further justifies the evidence-based factor in the policy making process. 6

Law has also been influenced by the evidence-based movement sweeping almost each and every area in natural and social sciences, which can be inferred by the explosive growth in empirical legal work. Although arguably empirical legal scholarship is not the same as

---


evidence-based law,\(^7\) the former enables relevant data collection and analysis processes that constitute the foundation of the latter. In this respect, there are two major advantages for law academics to actively participate in empirical research. The first is linked to the evolving and more accessible technology which facilitates academics in terms of access to data, making the foundational evidence-based work possible. The second is the increasing interdisciplinary collaborative opportunities among academics from different backgrounds, exposing law scholars to a more empirically-friendly research environment.\(^8\)

The importance of empiricism in legal research cannot be overlooked and much has been written about the virtues of evidence-based law.\(^9\) However, the real value of such empirical work lies more in its application rather than a focus at the level of academic discussion and exchange. To some extent, this difference distinguishes empiricism in law from evidence-based law. As pointed out by Prof. Rachlinski, ‘the point is to create better law – law informed by reality’.\(^10\)

The need to integrate more evidence in terms of law reform has been widely recognised by countries such as Australia. For example, the Australian Legal Reform Commission (ALRC) commented in the Managing Justice report\(^11\) that:

‘[The] depreciation of the legal system and failed efforts at reform often proceed on the basis of anecdote and assumption. This can include both untested and unfounded criticism of some current practices, procedures and institutions, as well as uncritical acceptance of alternatives.’

The ALRC has also noted that ‘law reform recommendations … need to be anchored in an appropriate evidence base’.\(^12\) Nevertheless, it is clear that the legal sector lags behind other disciplines in terms of the evidence-based movement and more work is required to meet this gap. A number of factors are responsible for the ‘law lag.’ In addition the law lag does not necessarily apply evenly across the justice sector. In the criminal area for example, a range of bodies continue to inform policy and funding.\(^13\) In the family law area, there may also be a greater focus on evidence based approaches and more reliance on empirical evidence.\(^14\) In the civil justice area however, despite some federal focus on this area\(^15\), it remains relatively uncommon for evidence based work to underpin academic work which may focus more on doctrinal developments.

First, the nature of law making itself may present an inherent obstacle to evidence-based law. This is partly because legislation is closely related to and influenced by the political process. This nature of law means that for legislators, during both the law making and reform

---

\(^7\) Rachlinski, above n 3, 910.


\(^10\) Rachlinski, above n 7.


\(^12\) Ibid.


processes, not only the benefits and adverse consequences of legislation suggested by the scientific research will be considered, as political gains will often be taken into account. As Antokolskaia notes, evidence-based legislation is developed when ‘the legislator in [their] choices for legislative interventions take[s] a rational and focused approach and does not let [themselves] be guided by JUST political and ideological reasoning, but also by relevant results of scientific inquiry assessing the effectiveness of those interventions.’ 16 It is therefore somewhat difficult to assess how much evidence the legislator would take into account and echoes Professor Rachlinski’s concerns about the prospect of empirically informed legal policy.17

Second, methodological concerns about poorly designed empirical research can undermine the value of evidence-based law. Empirical legal research, based on data collection and analysis, can involve surveys and interviews as well as interrogation of existing data sets. Issues can arise in the context of the design of appropriate methodologies in part because legal academics often have little training in terms of empirical methodologies. This can cause issues in terms of the capacity of legal researchers to interrogate data, or to design survey and interview questions in order to gain useful and reliable evidence. Poorly planned research, may trigger the question of ‘how evidence-based is this evidence-based research?’

Unlike the concerns expressed above about the nature of law and policy changes, this methodological issue could be addressed (at least to some extent) by providing more training to legal scholars and helping them develop skills in empirical research. 18 In addition, recognising and learning research methods adopted by colleagues in areas other than law, such as psychology and social science, is a feasible way for legal academics to facilitate their own empirical research and expand their contribution to evidence-based law. To some extent, this special issue of The Newcastle Law Review to include a couple of articles focusing on methodologies and we hope that our readers will benefit from this part of our peer-reviewed collection.

Arrangements to gather evidence in law can also be complex. Issues relating to confidentiality, consent and to some extent culture can impact on attempts by researchers to explore data and to undertake and extend research. For example, an inquiry about legal costs, might rely upon lawyers sharing their data about their costs. Such an undertaking would be dependent on lawyers being engaged in such research and to some extent being satisfied that such material would be kept confidential. More developed studies in the same area might require consideration of court data which may raise issues relating to stretched court budgets as well as interest in terms of research objectives. There may also be restrictions in terms of comparative research. This is particularly the case when research may be conducted in a foreign country and where the government has restrict regulations in place on research projects conducted by overseas researchers. 19

17 Rachlinski, above n 3, 922.
18 Bell, above n 8, 268.
This *Newcastle Law Review* issue is comprised of a combination of articles arising from the ‘Symposium on Evidence Based Law and Practice’ held by Newcastle Law School in May 2017. This one-day workshop attracted more than 50 participants who shared their latest research on evidence-based law and practice and generated some valuable interdisciplinary discussions on evidence-based approach among academics from law and other fields, including social workers. The issue includes six articles and one research note all with a focus on how evidence-based methods have been utilised by the academics and the practitioners in their research and practice, including family and children protection issues, reform on jury’s decision-making process, cyber security and open data policy and mediation.

We thank all the authors for their interest in contributing to and publishing with *The Newcastle Law Review*. Primarily though, we would like to show our great appreciation to The Hon Justice Margaret J. Beazley AO as we are very fortunate to have the permission to include Her Honour’s paper titled ‘Language: Law's Essential Tool’. This article was further developed by Her Honour on her speech at the 25th Sir Ninian Stephen Lecture at Newcastle in 2017. As law academics, we appreciate the vital importance of considering the language of law in terms of evidence-based approaches and Her Honour’s article assists to explain how the language has impacted the law, including evidence-based law. We also would like to thank all the colleagues who contributed to this issue.

**Professor Tania Sourdin**

Dean and Head of School, Newcastle Law School

**Dr Bin Li**

General Editor, *The Newcastle Law Review*
LANGUAGE: THE LAW’S ESSENTIAL TOOL

The Hon Justice Margaret J Beazley AO*

Introduction

1 The law is a wonderful and exciting profession of which Sir Ninian Stephen was the exemplar. A profession, by definition, carries with it the concept of service to others. There are, of course, many professions, some of them not so honourable. Indeed, there are those who suggest that law is one such profession. The old joke ‘how many lawyer jokes are there?’ elicits the response ‘only three – the rest are true stories’ whilst the question ‘how many lawyers does it take to change a lightbulb?’ is met with ‘three – one to climb the ladder, one to shake it and one to sue the ladder company’.

2 In keeping with this thrice times theme, I will first examine the role and impact that language has on law and legal concepts. Secondly, I will explore the way that concepts in the law have both remained constant whilst at the same time legal concepts and principles have developed through the use of language. Finally I wish to say something about the venerable jurist Sir Ninian Stephen, who passed away late last year, and the legacy he leaves us.

3 Before doing so, it is always salutary to remind ourselves that as lawyers, we play a central role in the administration of justice. Stated in slightly different terms, we are the custodians of the rule of law, an organising principle of our democratic society. Despite its ancient foothold in the laws of England, many of which themselves derive from the laws of Ancient Rome, the law continues to evolve, with sensitivity to today’s issues and in a manner that provides us with intellectual challenges and new horizons throughout our legal careers. At the outset, however, it is necessary to make the following disclaimer. This paper is not an excursus on legal history or culture per

* President of the New South Wales Court of Appeal and Adjunct Professor, University of Notre Dame, School of Law, Sydney. I wish to express my thanks to my Researcher, Brigid McManus, for her research and assistance in the preparation of this paper.
se, nor an examination of the principles of law but rather of the relationship between
language and the law.

4 It is self-evident that law comes into being and is given form by language. Sir Ninian
is praised for judgments that were clear and concise. One would hope that all legal
writing was thus. But like lawyers and the law itself, the language of the law often has
harsh critics. In *Bleak House*, Charles Dickens likened legal language to ‘street mud
which is made of nobody knows what … and when there is too much of it, we find it
necessary to shovel it away’.¹

5 One would have hoped that since those sharp words were written in the 1850s lawyers
might have cleaned up their act. However, in a letter written on 27 January 1972, one
solicitor wrote in the following terms to another solicitor in respect of a conveyancing
transaction:

Preparatory to the occurrence of completion herein, I furnish herewith for your
approval:

1. the instant Memorandum of Mortgage in duplicate signed by the
Mortagors and apposite witness consonant with your pertinent requisition.

...  

3. The Mortgagor’s authority respecting disposal of the advance noting
that at completion we shall by endorsement thereon confirm our oral request when
appointing completion herein.²

6 Language such as this is fodder to the sceptics who suggest that legal language is a
tool fashioned for obscuring the true meaning of statements, documents and legal
principle.³ The sceptics are correct in referring to language as a tool. But rather than
shrink from the barb, as lawyers, we should embrace it. Language is the tool by which
we negotiate and create contracts. It defines principle and enables us to persuade
others of our case. At a higher level of abstraction, language provides insight into the

¹ Charles Dickens, *Bleak House* (Bradbury and Evans, 1853) 153.
² Copy on file with author.
suggest that sentences in legal writing tend to be longer than in other disciplines and that their structure is
particularly complex: see Risto Hiltunen, ‘The Grammar and Structure of Legal Texts’ in Peter M Tiersma and
jurisprudential basis of the law and how the law works in practice. In these various ways, language, used appropriately, aids and strengthens the rule of law.

The Language of the Law

Any discussion of the language of the law calls for some preliminary observations about the rich history of legal language, some of which has its historical roots in Latin. Terms such as ratio decidendi\(^4\) and obiter dicta\(^5\) are undoubtedly familiar, as would be the nemo dat\(^6\) rule and the ex turpi causa\(^7\) principle. Recently the High Court reintroduced the archaic legal term felo de se into the legal lexicon.\(^8\) These terms, perhaps with the exception of the last, are well-known and the underlying principles which they expound are well-understood.

Although such Latin expressions are the source, in part at least, of the complaint that the law is inaccessible, we need not be overly apologetic about their use. We don’t require doctors to give up terms such as tibia and fibula. Fibula, incidentally, comes from the Latin word figo which means fasten and in the 17\(^{th}\) century a fibula was a buckle, brooch or clasp. Similarly botanists have yet to abandon the Latin names of plants. Take monstera deliciosa as an example, described in that scholarly online research facility Wikipedia as an arum, a genus of flowering plants within the Araceae family, which describes an epiphyte with aerial roots.\(^9\) I trust I have made my point!

The continued use of terms such as ratio and obiter reflects not only the deep historical roots of our law but the continuity of the legal concepts and principles they represent and thus their ongoing relevance. This is unsurprising in our common law

---

\(^4\) Meaning ‘the reason for deciding’.
\(^5\) Meaning ‘by the way’.
\(^6\) Meaning ‘nobody gives’. The full expression of the rule is nemo dat quod non habet, meaning ‘nobody gives what he does not have’.
\(^7\) Meaning ‘from a dishonourable cause’. The full expression of the principle is ‘from a dishonourable cause an action does not arise’.
\(^8\) IL v The Queen (2017) 91 ALJR 764, 769-72. Felo de se is described as a ‘peculiar species of felony, a felony committed on oneself’: William Blackstone, Commentaries on the Laws of England (1769) Book 4, 190.
tradition based on precedent, but in any event a sound case should be made before such well entrenched terms are abandoned.

The old and embedded traditions of the law can also be found in the continuing influence of the language used by the earliest common law lawyers. Take, for example, the Elizabethan jurist Sir Edward Coke, who is quoted as saying that ‘justice must have three qualities’, it must be free, full and speedy. This notion finds similar expression in s 56 of the Civil Procedure Act 2005 (NSW), which provides that the overriding purpose of the Act is to ‘facilitate the just, quick and cheap resolution of the real issues in the proceedings’.

Notwithstanding the simplicity of this language, the statutory injunction in s 56 is not an idle or catchy aphorism. It is central to basic but important procedural issues such as whether a party will be granted an adjournment or how many times a party will be permitted to amend a statement of claim or other legal process. The requirements of the section play a key role in ensuring a successful commercial environment, as the High Court explained in Aon Risk Services Australia Ltd v Australian National University. In that case, Justice Heydon, after observing that ‘commercial life depends on the timely and just payment of money’, stated that ‘the efficiency or inefficiency of the courts has a bearing on the health or sickness of commerce’.

The failure to deliver outcomes in accordance with the statutory injunction embedded in s 56 also affects how people feel and react when they are exposed to the legal system. This is an important consideration in maintaining respect for justice as administered by the courts. Allsop J (as his Honour then was) was alert to this in White v Overland when he observed: ‘[l]itigation is not a game. It is a costly and stressful, though necessary evil’. Section 56 and the observations made in Aon Risk Services and White v Overland are directed as much to the profession as they are to the administration of justice by the courts. As lawyers we not only have extraordinary

12 (2009) 239 CLR 175.
13 Ibid [137] (Heydon J).
privileges, we also have enormous responsibilities in ensuring that, in its everyday application, justice bears the hallmarks identified by Sir Edward Coke.

Linguistic Theory and Legal Language

13 Academics and judges alike have analysed and commented upon the use of language in the law, although it is fair to say that academics tend, appropriately, to focus on linguistic theory whilst judges are more concerned with language in a practical sense. However, the one is integral to the other.

14 The work of the British legal philosopher Professor HLA Hart suggests that there are four concepts that provide insight into the nature of language and how language can elucidate the nature of law: context; diversity; vagueness; and the performative use of language. It is the first three of these that I wish to discuss in this paper.

15 ‘Context’ is a simple, indeed, obvious, concept. It applies where the meaning of a word depends on the circumstances in which it is used.\(^\text{15}\) Consider the difference in the meaning of the word ‘prune’ depending on its use. Used as a noun, the word refers to a dried fruit. Used as a verb, it means ‘to trim’. We can only make sense of which meaning is intended by considering the word in its context. This does not only mean the words surrounding it in a sentence, which serve to characterise its grammatical meaning. It may be the geographical location in which the word is uttered or the identity of the speaker which provides the context.\(^\text{16}\) The point is that context provides insight into the intended meaning of the word used.

16 Understanding context is crucial for good lawyering and good judging. Perhaps one of the clearest examples of why context is important is when a word is used incorrectly. In the English case *Mannai Investment Co Ltd v Eagle Star Life Assurance Co Ltd*,\(^\text{17}\) Lord Hoffman observed that ‘[i]t is a matter of constant experience that people can

\(^{15}\) Endicott, above n 3, 947.

\(^{16}\) Unless of course it was a botanist speaking who could be using the word in either sense!

\(^{17}\) [1997] AC 749.
convey their meaning unambiguously although they have used the wrong words’.\textsuperscript{18}

His Lordship explained that:

We start with an assumption that people will use words and grammar in a conventional way but quite often it becomes obvious that, for one reason or another, they are not doing so and we adjust our interpretation of what they are saying accordingly. We do so in order to make sense of their utterance: so that the different parts of the sentence fit together in a coherent way and also to enable the sentence to fit the background of facts which plays an indispensable part in the way we interpret what anyone is saying. No one, for example, has any difficulty in understanding Mrs Malaprop. When she says ‘She is as obstinate as an allegory on the banks of the Nile,’ we reject the conventional or literal meaning of allegory as making nonsense of the sentence and substitute ‘aligator’ by using our background knowledge of the things likely to be found on the banks of the Nile and choosing one which sounds rather like ‘allegory’.\textsuperscript{19}

\textit{Mannai Investment Co Ltd v Eagle Star Life Assurance Co Ltd} concerned a tenant who had given notice to terminate two leases on 12 January 1995, pursuant to a clause which provided that the lease could be terminated ‘on the third anniversary of the term commencement date’. The leases had commenced on 13 January 1992 and the trial judge held that, on their true construction, the date on which they could be terminated was 13 January 1995.\textsuperscript{17}

The Court of Appeal held that a notice to terminate stated to take effect on 12 January could not operate to take effect on 13 January. A majority of the House of Lords rejected this approach and held that the notices, objectively construed and bearing in mind their context, left no doubt that the tenant wished to terminate the leases on 13 January. As a result, although the date of termination was wrongly described as 12 January, the notices were effective to terminate the leases. Hence, context was used to give meaning to the language of the lease.\textsuperscript{18}

It is not uncommon for counsel appearing in matters before the court to offer a word and suggest it be given a meaning that overlooks the context in which it appears. This may be a word found within legislation or a term within a legal instrument, such as a contract or even a word uttered by a party to proceedings. A particularly striking example of a word being given a meaning removed from the context in which it was spoken is found in a recent case heard before the Louisiana Supreme Court.\textsuperscript{19}

\textsuperscript{18} Ibid 774.

\textsuperscript{19} Ibid 774.
The case concerned a suspect in an interrogation, who told detectives to ‘just give me a lawyer dog’. The Louisiana Supreme Court ruled that the suspect was asking for a ‘lawyer dog’ and as the state of Louisiana had no canine attorneys, the police were not required to stop questioning the suspect. This ruling overlooked not only the colloquial meaning of the term ‘dog’ but also the meaning given to the term by the suspect’s full request. The suspect had said to the detectives: ‘[t]his is how I feel, if y’all think I did it, I know that I didn’t do it so why don’t you just give me a lawyer dog ’cause this is not what’s up.’

As this makes clear, failing to consider words within their context risks misconstruing them. If this occurs in the interpretation of a statutory provision or clause in a legal instrument, it risks giving the statute or instrument an effect it simply does not carry, and was not intended to carry. For example, terms in contracts must be construed in light of their commercial context. Similarly, if we misconstrue case law in this way, the precedential strength of the common law may be compromised. And if we remove words said by parties, whether in civil or criminal proceedings, from their context, we risk undermining the administration of justice. Had the suspect in the Louisiana case been reprimanded for calling the police officer a ‘dog’, but provided with a lawyer, the administration of justice would have been rightfully served.

The second concept, diversity, is linked to the context principle. It recognises that variations in context can extend the application of a word in diverse ways. A word may have a single, clear application while its meaning varies in different contexts. For example, we are all familiar with adopted children referring to their adoptive parents as ‘Mum’ and ‘Dad’, although if one heard those terms being used without knowing the context, it is likely to be assumed that those being referred to were the child’s biological parents.

The meaning that a word may have in the legal context may vary from the meaning commonly understood by the lay person, or may vary between different areas of law.

---


21 See Codelfa Construction Pty Ltd v State Rail Authority of New South Wales (1982) 149 CLR 337; Mainteck Services Pty Ltd v Stein Heurtey SA (2014) 89 NSWLR 633.
and application. Returning to the example of the adopted child, the word ‘parent’ usually refers to a child’s biological parents but in some areas of law may refer to anyone charged with the care of a child.22 One can even go so far as to say that the Court in the exercise of its parens patriae jurisdiction performs a parental role.

24 Indeed, the word ‘Judge’ might fall into Professor Hart’s diversity category. I was once engaged in the following correspondence with a would-be author of a miscellany at law who was looking for legal anecdotes. It was addressed to ‘Justice’ and commenced ‘Dear Sir’. I responded as follows:

Dear Mr X,

Thank you for your letter … beginning ‘Dear Sir’. Perhaps your first anecdote could come from you. My first name is Margaret.

Not defeated, the would-be author responded. His letter commenced ‘OOPS’ and continued that he was embarrassed and would consider including the gaffe in the foreword of his book. He beseeched me again to provide him with anecdotes saying ‘perhaps something arises related to Her Honour’s femininity’. So far as I am aware, the miscellany never came into existence.

25 Thirdly, although we search for certainty in the law,23 Professor Hart points to aspects of the law where certainty gives way to what he describes as ‘vagueness’ in the sense that a word or circumstance may not be prescriptive or may not reflect a single given meaning or lead to a specific outcome.

26 Professor Hart uses the discretionary exercise of a power to explain ‘vagueness’. For example, whilst the exercise of a judicial discretion permits a range of outcomes, so that its normative expression is ‘non-prescriptive’, its freewheeling exercise is not permitted. It must be exercised judicially.24 As Professor Hart explains:

---

22 Endicott, above n 3, 950-951.

23 For example, a fundamental principle in the formation of a contract is that of certainty. In cases were the language of a contract is uncertain such that a court is unable to give the parties’ language a sufficiently clear and precise meaning in order to identify the rights and obligations agreed upon, there will be no concluded agreement. See, eg, G Scammell & Nephew Ltd v Ouston [1941] AC 251.

24 See, eg, Oshlack v Richmond River Council (1998) 193 CLR 72 in which Gaudron and Gummow JJ held, at 81, that the conferral of a discretion as to costs was to be ‘exercised judicially, that is to say not arbitrarily,
In every legal system a large and important field is left open for the exercise of discretion by courts and other officials in rendering initially vague standards determinate ….

27 An example is found in the use of the word ‘may’. When used in statutes and regulations ‘may’ will not always have the permissive meaning it otherwise bears. For example, in Finance Facilities Pty Limited v Commissioner of Taxation the High Court considered s 46(3) of the Income Tax Assessment Act 1936 (Cth) (since repealed) which provided that the ‘the Commissioner may allow’ a taxpayer a further rebate if satisfied of the matters specified in the section.

28 A majority of the Court held that despite the permissive nature of the words ‘may allow’, if the Commissioner was satisfied of the matters set out in the section, the Commissioner was obliged to allow a further rebate. Windeyer J explained that as the scope of the permission or power given was circumscribed, this was one of those cases in which ‘the “may” becomes a “must”’. The use of the word ‘shall’ provides another example. As Earl Cairns LC explained in Julius v Bishop of Oxford:

The words ‘it shall be lawful’ are not equivocal … They confer a faculty or power, and they do not of themselves do more … But there may be something in the nature of the thing empowered to be done, something in the object for which it is to be done, something in the conditions under which it is to be done, something in the title of the person or persons for whose benefit the power is to be exercised, which may couple the power with a duty …

26 But see, eg, Acts Interpretation Act 1901 (Cth), s 33(2A); Interpretation Act 1987 (NSW), s 9. Similar provisions have been enacted in other Australian jurisdictions.
27 Section 46(3) provided that:

Subject to the succeeding provisions of this section, the Commissioner may allow a shareholder, being a company that is a private company in relation to the year of income and is a resident, a further rebate … if the Commissioner is satisfied that

(a) the shareholder has not paid, and will not pay a dividend during the period commencing at the beginning of the year of income of the shareholder and ending at the expiration of ten months after that year of income to another private company …

29 (1880) 5 App Cas 214, 222-3.
Interpreting Legal Language

30 An understanding of Professor Hart’s linguistic concepts offers two primary insights into the practice and development of the law. The first, and perhaps the area in which the study of language and linguistics has had greatest impact, concerns legal interpretation.30 In the area of constitutional law one only need consider the debate between originalists and those who propound a ‘living tree’ approach to constitutional interpretation to appreciate how vague, uncertain and context-dependent constitutional language can be.31

31 In the area of statutory interpretation, consider, for example, the Acts Interpretation Act 1901 (Cth) which provides in s 15AA that when interpreting legislation, the interpretation which would ‘best achieve the purpose or object of the Act’ is to be preferred. Similarly, the Interpretation Act 1987 (NSW) provides in s 33 that the ‘construction that would promote the purpose or object underlying the Act or statutory rule’ is to be preferred. Both Acts provide that a Court may have regard to extrinsic material in interpreting a provision where that provision is ‘ambiguous or obscure’.32

32 This broadly reflects the common law approach to interpretation known as the ‘purposive approach’. The purposive approach looks beyond the literal meaning of the words where the literal meaning of the text is ambiguous, so as to interpret them in light of the underlying purpose of the Act or the ‘mischief’ the provision seeks to address.33 The Interpretation Act provisions can also be understood as part of what Chief Justice Spigelman of the Supreme Court of New South Wales described in 2007 as a shift from ‘text to context’ and from ‘textualism to contextualism’ in constitutional, statutory and contractual interpretation.34

33 The Interpretation Act provisions, and the principles of statutory interpretation recently articulated by the High Court, emphasise the balance in interpretative method

---

30 Michael Freeman and Fiona Smith, ‘Law and Language: An Introduction’ in Michael Freeman and Fiona Smith (eds), Law and Language (Oxford University Press, 2013) 1, 5.
31 See eg, the conflicting approaches taken by the US Supreme Court in District of Columbia v Heller 554 US 570 (2008) with respect to the Second Amendment right to bear arms.
32 Acts Interpretation Act 1901 (Cth) s 15AB(1)(b)(i); Interpretation Act 1987 (NSW) s 34(1)(b)(i).
between text and context. This is evident in *Alcan (NT) Alumina Pty Ltd v Commissioner of Territory Revenue (Northern Territory)*, where Hayne, Heydon, Crennan and Kiefel JJ explained the principles of statutory construction in terms with which we are well familiar:

> the task of statutory construction must begin with a consideration of the text itself. The meaning of the text may require consideration of the context, which includes the general purpose and policy of a provision, in particular the mischief it is seeking to remedy.

34 Similarly, French CJ explained that statutory interpretation involves looking to ‘the ordinary and grammatical sense of the statutory words to be interpreted having regard to their context and the legislative purpose’ (emphasis added).

35 The place of context in statutory construction was again examined in *Certain Lloyd’s Underwriters v Cross*, where Kiefel J (as her Honour then was) explained that:

> The starting point for [the] process of [statutory] construction is the words of the provision in question read in the context of the statute.

36 This integrated approach to text in context was reiterated more recently in *SZTAL v Minister of Immigration and Border Protection*, where Kiefel CJ, Nettle and Gordon JJ stated that that:

> The starting point for the ascertainment of the meaning of a statutory provision is the text of the statute whilst, at the same time, regard is had to its context and purpose. Context should be regarded at this first stage and not at some later stage and it should be regarded in its widest sense.

37 This recognition of the significance of context and purpose may be seen as involving a recognition of Hart’s context and diversity principles, although not expressed in precisely those terms.

---

36 (2009) 239 CLR 27, [47].
37 Ibid [4].
38 (2012) 248 CLR 378, 412. Her Honour continued: ‘[c]ontext is also spoken of in a broader sense as including the general purpose and policy of the legislation, in particular the mischief to which the statute is directed and which the legislature intended to remedy.’
39 (2017) 91 ALJR 936, [14].
The problems which can arise when context and purpose, or context and diversity, in Hart’s terminology, are overlooked and a strictly literal interpretive approach is adopted are starkly apparent in the decision of the Full Court of the Supreme Court of Western Australia in *Higgon v O’Dea*.\(^{40}\) In that case, the Court considered the effect of s 84 of the *Police Act 1892* (WA), which provided that:

> Every person who shall have or keep any house, shop, or room, or any place of public resort, and who shall wilfully and knowingly permit drunkenness or other disorderly conduct … or knowingly suffer any unlawful games or any gaming whatsoever therein, or knowingly permit or suffer persons apparently under the age of sixteen years to enter and remain therein … shall … be liable to a penalty of not more than five pounds.

The defendant owned an amusement arcade and was charged under the Act with permitting persons under the age of 16 to enter. The arcade did not permit gaming or disorderly conduct. The Court held that despite this, and the fact that the provision operated ‘absurdly, unjustly and unreasonably’\(^ {41}\) – for example, under a literal reading no children would be permitted in shops – its words were clear and the defendant had committed an offence.

Even the language used in the provisions of the *Acts Interpretation Act 2001* (Cth) and the *Interpretation Act 1987* (NSW) relating to the use of extrinsic materials is worthy of note. These provisions allow for consideration of extrinsic materials where the language of a statute is ‘ambiguous or obscure’. In *R v Sharma*,\(^ {42}\) Spigelman CJ noted that the use of the both of these two words seemed curious but considered that it may be explained by the distinction drawn between them by a member of the House of Lords in the English case *Ellerman Lines Ltd v Murray*.\(^ {43}\)

In *Ellerman Lines Ltd v Murray*, their Lordships were unanimous that a statute which provided for the payment of wages for a period of two months to sailors whose employment was terminated as a result of their ship being wrecked or lost and where they remained unemployed, was unambiguous. However, two law lords held that payment must be made irrespective of whether the sailor would have been employed but for the wreck or loss of the boat. A further two law lords considered that payment

\(^{40}\) [1962] WAR 140.
\(^{41}\) Ibid 142.
\(^{42}\) (2002) 54 NSWLR 300, [54].
\(^{43}\) [1931] AC 126.
must be made unless the owner of the ship could show that the sailor would not have
been employed. One law lord, Lord Blanesburgh, held that payment need only be
made for the sailor’s contracted period of employment.

Lord Blanesburgh considered the distinction between ambiguous and obscure
language in a passage which is worth repeating. He began by noting that the section
did not bear its meaning ‘upon its sleeve’ and continued by observing that:

It yields up its secret only to the patient inquirer; its truth lies at the bottom of the
well. It is obscure, it remains oblique, but it is not in the result ambiguous. The truth
from the well is found, at the end of the search for it, to have been leaking out of the
section itself all the time just as the truth … may leak out sometimes even from an
affidavit.44

Recognition of the importance of context can also be found in the jurisprudence
concerning the construction of contracts. In Mainteck Services Pty Ltd v Stein Heurtey
SA, the New South Wales Court of Appeal emphasised that:

to say a legal text is ‘clear’ reflects the outcome of [the] process of interpretation. It
means that there is nothing in the context which detracts from the ordinary literal
meaning. It cannot mean that context can be put to one side …45

Similarly in the United Kingdom, in a case concerning patents, Lord Hoffman has
noted that:

No one has ever made an acontextual statement. There is always some context to any
utterance, however meagre.46

An understanding of the role played by context in shaping meaning has resulted in
less reliance on dictionary definitions, although this is far from a modern
phenomenon. In Mainteck Services Pty Ltd v Stein Heurtey SA, Leeming JA quoted
from Lord Herschell, who in 1898, said that the words in a clause of a contract must
not be given a meaning that encompasses ‘everything that might be said to come

44 Ibid 144.
45 (2014) 89 NSWLR 633, [77] (Leeming JA).
46 Kirin-Amgen Inc v Hoechst Marion Roussel Ltd [2005] 1 All ER 667 [64].
within a possible dictionary definition use of them’. Rather, the words ‘must be interpreted in a way in which business men would interpret them’.  

A century later, Mahoney JA expressed a reluctance to rely on dictionary definitions on the basis that ‘[d]ictionaries are not a substitute for the judicial determination of the interpretation and then construction of statutes and other documents’. This is because

> the meaning of the words used in a statute or document is not merely the sum of the individual meanings of the words used, ascertained from dictionaries … a word is the skin of a living thought and it is the thought which the court must ascertain and apply.

### The Development and Designation of Legal Terms

The second way in which the study of language and linguistics offers insight into the law concerns the development and designation of legal terms. Indeed, Socrates is said to have counselled that wisdom begins with the definition of terms. Legal terms can be understood as the names or labels given to an underlying legal concept or principle. Reference has already been made to Latin terms such as *ratio decidendi* and the *nemo dat* rule. As words, they tell us little. As words which relate to, describe, or define a legal concept or principle they convey a deep and precise meaning. It could be said that such terms are the external expression of the underlying legal concept or principle.

In this respect, a distinction needs to be drawn with purely descriptive terms. We are all familiar with the Shakespearean aphorism that a ‘rose by any other name would smell as sweet’. However, this is not true in law. Legal terms, properly used, are inextricably tied to the underlying legal concept. It would not do, for example, to call the legal concept embodied in the term *ratio decidendi* the *nemo dat* rule.

---

48 Ibid 444.
50 Ibid 560.
52 William Shakespeare, *Romeo and Juliet*, act II, scene II.
Accordingly, legal scholars recognise that the creation and development of useful terminological tools is ‘one of the central challenges of the discipline’.\textsuperscript{53}

**Linguistic Choices and their Consequences**

Recognition of the significance of legal terms leads to an appreciation of the consequences of our linguistic choices and the manner in which words convey subtle but significant messages. An example can be found in the standard of the ‘reasonable man’ and the surrounding critique of the concept. The standard of the ‘reasonable man’ – the man on the Clapham omnibus or the Bondi tram – is used as an objective standard, a stand-in measure for what a reasonable person might think or do in the circumstances.

Traditionally the language of the standard was explicitly gendered. In the 1980s, feminist critiques of the test led to a shift in this language – the ‘reasonable man’ became the ‘reasonable person’. This change was not only appropriately inclusive, but recognised that, implicitly, the notion of the ‘reasonable man’ gave a degree of precedence to male experience.\textsuperscript{54} In this respect, the change was part of a broader shift toward gender-neutral legal language, which also saw general references to ‘he’ or ‘his’ replaced with ‘he or she’ and ‘his or her’ and, more recently, the grammatically incorrect ‘them’ and ‘their’ when referring to the singular. Does such language make a difference? Anecdotally and from personal experience it most certainly does.

I recall one memorable incident that occurred in my early days sitting as a judge in the Court of Appeal in which a barrister, using the word ‘draftsperson’ to refer to the drafter of a badly worded contract, was continually corrected by the presiding judge to use ‘draftsman’. The barrister would look to me with sheer fear on his face, back at the presiding judge with even greater trepidation and mumble something which sounded like a cross between the two. Eventually, I put an end to the exchange by simply telling the barrister that it was alright to use the word ‘draftsman’ as a ‘draftswoman would never have drafted such a bad contract’.

\textsuperscript{53} Endicott, above n 3, 938.

While this is a humorous story, it does have a more serious undertone, in that it demonstrates a degree of resistance to changing the language of the law in response to changing social values and norms. Not even the rich history of legal language can justify such resistance. To resist, in such cases, reflects stultification rather than the vibrancy which a living law requires. In other words, it has nothing to do with language or law and much to do with obstinacy – a last stand as it were, the resisters clinging to the comforts of the ‘old world’.

Returning to the notion of the ‘reasonable person’, concerns have been raised that the simple substitution of the word ‘person’ for ‘man’ is not enough to ensure that the underlying concept is appropriately applied to all members of the community. On this view, a simple change to gender-neutral phraseology is not always sufficient to ensure gender-neutral application. A change in language can be an apt starting point but it must be accompanied by changes in the way the law is applied to recognise a broader range of experiences.

Recently, the language used to discuss domestic violence has been the subject of attention, both academically and more broadly in the media. Here, the concern is that ‘words can convey assumptions or obscure stereotypes using a veneer of objectivity’. In this respect, certain language can ‘perpetuate violence, and silence can be used to render issues and persons invisible and their experiences “unutterable”’. For instance, it is not unusual to hear it said that an offender ‘lost’ his or her ‘temper’ or ‘snapped’ or was motivated by ‘jealousy’ or ‘anger’. Such terms may not immediately sound problematic but they can be used to suggest some blameworthy or inflammatory conduct on the part of the victim which, in turn, prompted a loss of control on the part of the offender. It has been suggested that:

57 Ibid 50.
58 Ibid 50.
These discursive strategies when used by the perpetrator or defence appear designed to ameliorate the perpetrator’s responsibility for the crime, and are reinforced by the court’s acceptance and replication of an offender’s explanations.59

The focus on a loss of control can therefore obscure the reality of the dynamics of domestic violence. Rather than framing an offender’s behaviour in terms of a loss of control, an offender’s actions may, in a given situation, be more correctly framed as an attempt by the offender to maintain control over the victim.60 The use of appropriate language to accurately characterise what occurred will more aptly capture the degree of wrongdoing and thus be more likely to lead to appropriate remedies or punishment.

Similar affects can be observed in the description of a relationship as ‘turbulent’, ‘rocky’, ‘volatile’ or ‘stormy’. Once again, this language is problematic in that it removes agency from both the offender and the victim and subtly attributes it to the ‘relationship’.61

Just as gender-based scholarship offers insight into the significance of the terms used to describe legal concepts and the assumptions underlying them, so too does critical race theory. This is evident in discussions regarding the regulation of language, particularly in the context of racial vilification – an area where the balancing of prohibitions with free speech has proved both problematic and divisive.

Scholars of critical race theory note that racially vilifying language is often framed in terms of ‘causing offense’ and argue that when such language is described as merely ‘offensive’, it undermines the nature of the harm inflicted. They note that ‘offensive’ is used ‘as if we were speaking of a difference in taste’, overlooking the fact that so-called offensive language can cause tangible injury:

There is a great difference between the offensiveness of words that you would rather not hear because they are labeled dirty, impolite, or personally demeaning and the injury inflicted by words that remind the world that you are fair game for physical

59 Ibid 51.
60 Ibid 51.
61 Ibid 52.
attack, that evoke in you all of the millions of cultural lessons regarding your inferiority that you have so painstakingly repressed ...  

The harm that can be experienced as a result of such language has been reported as having both short- and long-term effects and includes physiological symptoms and emotional distress ranging from an increased pulse rate and difficulty breathing, to nightmares, post-traumatic stress disorder, hypertension and, in the most severe cases, self-harm. Racially vilifying language may also have an effect on one’s sense of self-esteem and personal security and can result in feelings of being silenced, both individually and as a group, and being excluded from the broader community. Understanding that language considered offensive can cause harm of this nature, as opposed to simply causing offence, has very real consequences for the manner in which we balance prohibitions on racially vilifying language with free speech.

What I trust these few examples demonstrate is that language shapes thought and the precise use of language and careful choice of the words we use is important. The extent to which language shapes thought, which in turn shapes law, is, I suggest, becoming evident in Victorian jurisprudence following the introduction of the Charter of Human Rights and Responsibilities Act 2006 (Vic). The Charter is only the second of its kind in Australia, where rights have traditionally been protected by the common law. One of the driving forces behind its introduction was a desire to make rights protections clearer and more transparent and to promote a rights-based dialogue within parliament, the courts and the public more broadly.

By fostering an explicit language of rights, the Charter can be observed to have encouraged a shift in the manner in which judges approach other areas of law and the issues that arise within them. This is apparent, for example, in the case of Boulton v

R, a Victorian Court of Appeal guideline judgment which considered the sentencing option of Community Correction Orders (‘CCOs’). CCOs are flexible, non-custodial orders to which coercive and rehabilitative conditions can be attached.

The Court (Maxwell P, Nettle, Neave, Redlich and Osborn JJA) listed a number of ways in which a CCO is punitive as opposed to merely rehabilitative. In assessing the punitive effect of a CCO, the Court stated that:

The (relative) severity of a penal sanction can be assessed by reference to its impact on the offender’s rights and interests. The more important the rights and interests intruded upon, and the more significant the intrusion, the severer is the sanction. Attention should therefore be directed to the degree to which the sanction will affect fundamental rights and interests such as the offender’s freedom of movement, choice regarding his/her activities, choice of associates, and privacy. (citations omitted)

The language of fundamental rights and interests seen here, whilst not unknown in the common law, is redolent of human rights discourse.

**Rule of Law and Legal Language**

Experience has demonstrated that the interplay between language and the law can facilitate and strengthen the rule of law. There are many formulations of the rule of law. One of the earliest descriptions is that of Sir Edward Coke, who declared that ‘the King should not be under man, but under God and the law’. This is, in effect, a pronouncement of government under law, the idea that the monarch and his council should act through and under the law, and not through the exercise of prerogative powers. It reflected the concerns of the time, a period in which there was growing tension regarding the powers of the monarch and the scope of those powers.

One of the most influential enunciations of the rule of law is that of the English jurist A V Dicey, who is credited with popularising the term. Dicey’s conception involved ‘at least three distinct though kindred conceptions’. First, that ‘no man is punishable

---

66 (2014) 46 VR 308.
67 Ibid [90].
69 *Prohibitions del Roy* (1607) 12 Co Rep 63, 65.
or can be lawfully made to suffer in body or goods except for a distinct breach of law established in the ordinary legal manner before the ordinary courts of the land.\textsuperscript{73} Secondly, that ‘no man is above the law, but … every man, whatever be his rank or condition, is subject to the ordinary law of the realm and amenable to the jurisdiction of the ordinary tribunals’.\textsuperscript{74} And thirdly, that ‘the general principles of the constitution … are … the result of judicial decisions determining the rights of private persons in particular cases brought before the courts’.\textsuperscript{75}

This formulation has been criticised for placing too great a focus on the form the law takes and failing to consider its substance. This focus has very real practical consequences. One only need look at the experience of Germany during the Third Reich or South Africa during Apartheid to see this. Both these regimes could be said to have complied with a Diceyan formulation of the rule of law but could hardly be said to be fair and just societies.

A more recent formulation is the Rule of Law Index, which was developed by the World Justice Project and is used to measure the state of the rule of law in 97 different countries.\textsuperscript{76} The index looks to features including accountability under the law generally applicable to governments, public officials, individuals, and public and private entities; clear, publicised, stable and just laws evenly applied which protect fundamental rights including the security of persons and property; accessible, fair and efficient processes for the enactment, administration and enforcement of laws; and timely delivery of justice by a sufficient number of competent, ethical, independent, adequately resourced representatives and neutrals who reflect the makeup of the communities they serve.

Sir Ninian himself offered a formulation of the rule of law as ‘not one simple ideal but rather a group of vital principles’.\textsuperscript{77} Like the biblical Ten Commandments, he described these principles as negative in nature, ‘descriptive of what should not occur’

\textsuperscript{73} Ibid 188.
\textsuperscript{74} Ibid 193.
\textsuperscript{75} Ibid 202.
\textsuperscript{77} Sir Ninian Stephen, ‘The Rule of Law’ (2003) 22(2) \textit{Academy of Social Sciences} 8, 8.
and ‘what should not be done’. Sir Ninian then suggested that the rule of law embodied four ‘cardinal’ principles. First, ‘that government should be under law, that the law should apply to and be observed by Government and its agencies, those given power in the community, just as it applies to the ordinary citizen’. Secondly, ‘that those who play their part in administering the law … should be independent of and uninfluenced by Government in their respective roles’. Thirdly, ‘that there should be ready access to the courts of law for those who seek legal remedy and relief’. And fourthly, ‘that the law of the land should be certain, general and equal in its operation’.

The common thread that wends through these various formulations is that of government under law. However, the language through which the rule of law is now expressed encompasses broader notions responsive to modern times. The addition of requirements such as ‘access to the courts’, ‘equal’ application, or ‘just’ laws gives the concept an enriched meaning in that it demonstrates its adaption to changing social contexts.

A notable feature of more recent formulations of the rule of law is the requirement that the law be ‘clear’ and ‘certain’. This depends significantly on the language through which it is expressed and the language used by legal professionals and the courts. One approach to ensuring clarity and certainty is to use ‘plain language’. While simple and direct expression has long been a focus in some corners of the legal community, the plain language movement first gained strength in the 1970s as a result of various initiatives which centred on the belief that the public at large should be able to understand their rights and obligations. The plain language movement eschews traditional legal phraseology and expression in favour of producing:

language and design that presents information to its intended readers in a way that allows them, with as little effort as the complexity of the subject permits, to understand the writer’s meaning.

---

78 Ibid 8.
80 Ibid 68.
Plain language is one approach to achieving clarity and certainty but it is not the only means by which to ensure that legal language clearly and accurately reflects underlying legal concepts in a manner that strengthens the rule of law. While plain language is desirable, it is sometimes necessary, in order to capture the rich complexity that makes law so effective, to use language directed to more precisely capturing that complexity.

This is certainly not to say that we need not consider the language we use. In fact, quite the opposite – language is important and precise language the goal. An understanding of language and how it works assists in construing legal texts such as statutes and contracts, aiding our understanding of law and how it works. An understanding of language also allows us to recognise the significance and consequences of the legal terms we use, as well as our word choice more generally. As these two points illustrate, precise language assists in developing clear and certain law. Precise language, which may be plain language – but is not necessarily – thus plays a key role in building and strengthening the rule of law.

Conclusion

If we are looking for the model use of precise language, we cannot go past Sir Ninian Stephen. Sir Ninian is widely praised for ‘his enviable communication skills – his lucid writing style, his compelling turn of phrase and his beautiful voice’. It certainly does not stop there. Sir Anthony Mason observed in Sir Ninian’s judgments ‘an elegance of literary style, a lightness of touch, indeed an elusive quality’. Sir Ninian’s judgments were ‘easy to read, a world apart from the dense, grinding judicial style which is characteristic of typical High Court judgments’. His influence has been enduring.

---


83 Ibid 5.

84 For example, the High Court continues to undertake the task of constitutional characterisation in accordance with his approach in Actors & Announcers Equity Association of Australia v Fontana Films Pty Ltd (1982) 150 CLR 169.
Sir Ninian’s career, and his influence, is not limited to his time on the High Court. In fact, he left the Court prematurely to take up an appointment as Governor-General of Australia. In this position, he displayed the same deftness, impartiality and commitment to public service that he demonstrated as a judge, earning him praise as ‘the very model of a modern Governor-General’. His time as Governor-General was certainly no quiet retirement – over the course of his appointment he is said to have delivered nearly one thousand speeches. Unsurprisingly then, his words continue to be his most enduring legacy. For example, his description of the governor-general’s role as being to ‘represent … the Australian nation to the people of Australia’ is oft-repeated and said to remain ‘one of the best “job descriptions” of the vice-regal office’.

Following his term in office, Sir Ninian was appointed Ambassador for the Environment in 1989, one of the first such positions anywhere in the world. Sir Ninian went on to serve as an International Peace Envoy in South Africa, Northern Ireland, Bangladesh and Burma. He also served at the Permanent Court of Arbitration, the International Court of Justice and the International Criminal Tribunal for the former Yugoslavia, as well as serving an investigatory function in relation to the Khmer Rouge atrocities in Cambodia.

This string of extraordinary appointments demonstrates Sir Ninian’s unyielding commitment to public service. They are also a testament to his skills as a diplomat, his political sensitivities and ultimately, his gift for clear, measured and graceful communication.

---

86 Ibid 31.
90 Ibid 119.
Evidence-based law and practice with disadvantaged populations:
Perspectives of parents involved in child protection proceedings+

Nicola Ross*, Lou Johnston**, Jessica Cocks***, Lynette Stoker****

Abstract: Lawyers and other professionals who work with disadvantaged populations often have limited research evidence about the impact of laws on their clients. People who experience socio-economic disadvantage can find it very difficult to access legal systems and processes. Locating and engaging them in research about their experiences is also challenging. One such group is parents who have children removed from their care due to child protection concerns. This article examines some of the prior research carried out with this group and then describes a qualitative study conducted with 18 parents in the Hunter Valley, New South Wales. This research explored their perspectives on the Children’s Court, child protection, and out-of-home care proceedings and processes. The findings suggest that parents whose children are removed are systematically disempowered. This raises major social justice and economic questions for government, policy-makers, practitioners, and researchers. The research demonstrates one approach to research with disadvantaged groups, which includes a multidisciplinary research team and assistance from parent consultants.

Key Words: evidence-based law, disadvantaged populations, child protection

* This is an extended version of a paper presented by Dr Nicola Ross and Lou Johnston at the Symposium on Evidence-Based Law and Practice, held in Newcastle on 22 May 2017, hosted by the Newcastle Law School, University of Newcastle. Jessica Cocks and Lynette Stoker are the other research team members. The team was assisted by two parent consultants, Teegan Bain and Felicity Kime, who had children removed and restored. Nicola and Lou wish to thank Ryan McGowan who provided able research assistance in developing the article in its early stages. The authors also wish to thank the anonymous reviewer of this article for their helpful comments.

** Senior Lecturer and Joint Theme Leader: Health, Justice & Social Affairs, Evidence-based Law and Practice Priority Research Initiative, Newcastle Law School, Faculty of Business and Law, University of Newcastle. Author contact: nicola.ross@newcastle.edu.au.

*** Practice Lead, Children, Families and Young People, Life Without Barriers and Conjoint Lecturer, Social Work, School of Humanities and Social Science, University of Newcastle. Author contact: lou.johnston@newcastle.edu.au.

**** Lecturer, Family Action Centre and Social Work, University of Newcastle. Author contact: jessica.cocks@newcastle.edu.au.

***** Lecturer, Family Action Centre and Social Work, University of Newcastle. Author contact: lynette.stoker@newcastle.edu.au.
Introduction
Parents who have their children removed experience numerous personal and social disadvantages. They are stigmatised in a similar way to those accused of criminal offences but have considerably less procedural protections against the actions of the state. To improve legal and professional practice with this disadvantaged group, it is important to understand how they experience legal proceedings and processes, and how their participation can be better supported. Lawyers and other professionals can find it challenging to acknowledge the vulnerabilities of parents who have abused or neglected their children, because of a perception that parents’ needs are in competition with children’s needs. While children’s safety and wellbeing should remain the paramount focus of proceedings, parents’ experiences and support needs are of research interest because parents and families are key to optimal outcomes for children.

This article first delineates why these parents are considered to be a disadvantaged group, then outlines prior research carried out with them. The methods that previous studies used and the available evidence about parents’ experiences of child protection systems and proceedings are examined. Secondly, the design and context of the research with parents in the Hunter Valley is described, and the primary themes and findings that relate to legal practice are presented. These are illustrated by examples of parents’ narratives taken from interviews and focus groups. The article concludes with suggestions on how this research might influence law, policy and practice.

Parent disadvantages in child protection proceedings
Most parents of children who are removed experience poverty and intergenerational disadvantage,¹ often compounded by substance use,² violence, disability, mental illness and social instability. Those factors affect a broad spectrum of their parenting experiences, including parenting capacity and child maltreatment,³ their children’s removal, participation in

² The statutory child protection agency in NSW, Family and Community Services (FACS), estimates that 80 per cent of all child abuse reports to Community Services involve parents experiencing substance use issues. See Department of Family and Community Services, Government of New South Wales, Our Services, Community services, <http://www.community.nsw.gov.au/about-us/our-services>.
court, child protection and out-of-home care processes, inclusion in their children’s lives while in care, and the likelihood of children being restored to their care. As such, disadvantages already exist prior to involvement in child protection proceedings and interventions. Subsequently, those disadvantages can be exacerbated, becoming more complicated, compounded, and cyclical during and after processes and interactions with different systems. Poverty can increase other adversities, such as,

increased conflict and stress, family instability, and neighborhood disorder…

[which] may in turn inhibit parenting capacity or negatively affect parenting practices through increased stress or decreased support… [that] frequently co-occur with child welfare involvement.4

Parents experience child removal and placement as stigmatising, isolating, disrespectful and unhelpful,5 and as involving processes which force them to comply with unrealistic targets,6 where they are locked out of participation in their children’s lives,7 and attacked and belittled.8 Participants in the Harries study talked about the increased isolation and disruptive and corrosive impact of statutory intervention in their lives and relationships — with each other, within and between families, within the community, and with helping services.9 This was echoed in Hinton’s study, where parents received little information from authorities about child protection processes, the progress of their matter, or what was happening to their children in out-of-home care. Those first contacts with child protection processes led parents to feel worthless, disrespected and powerless.10 Studies with Aboriginal parents revealed they did not feel respected by child protection authorities, who talked down to them rather than listening to them. A perceived lack of respect was in many instances based on the perception that workers

---

4 Fong, above n 1, 6.
9 Harries, above n 5, 25.
10 Hinton, above n 5, 43.
had treated parents or carers as if they were untrustworthy.\textsuperscript{11} Parents’ requests for help were rarely heeded: instead they led to the removal of their children. Stories of betrayal, broken promises, blame and shame, loss and grief, together with a lack of common courtesy and respect for others resonate through these studies.\textsuperscript{12}

Parents’ relationships with other people, including lawyers, workers, and carers, are key to their participation in proceedings and processes. The processes through which child protection authorities investigate and bring proceedings to court are complex. The stigma, fear and grief parents experience along with their lack of knowledge and social disadvantage mean they cannot engage with child protection processes effectively in the absence of trusting relationships. Research indicates that parents are less likely to engage actively and positively with these people if they feel judged, not listened to, or their parental role is undermined.\textsuperscript{13} Relationships with child protection professionals are often beset by poor communication, including a failure to listen, consult and to provide clarity,\textsuperscript{14} and a lack of sensitivity, understanding and empathy, which some parents experience as harassment.\textsuperscript{15}

\textbf{Prior research: methodologies}

Prior Australian studies which have investigated parents’ perspectives of the child protection system and out-of-home care have used primarily qualitative research methodologies. Common forms of qualitative research include interviews, focus groups, and ethnographic studies such as participant observation and life history interviewing. These methods are well-suited for research with disadvantaged groups, as they can provide an in-depth and detailed

\begin{footnotesize}
\begin{enumerate}
\item Klease, above n 5, 26.
\end{enumerate}
\end{footnotesize}
understanding of the experiences of people in their everyday life. Qualitative studies allow researchers to gain insights into the lives and perspectives of these parents, by engaging them in conversational interviews and other qualitative techniques. The development of reciprocal relationships between researchers and participants who are members of disadvantaged groups can also lead to a better exploration of perspectives and experiences. These methodologies generally involve conducting research with participants, rather than on participants.

The Family Inclusion Network’s 2005 study involved focus groups with parents to explore their experiences and perspectives. The focus groups were held across Queensland and involved 67 parents and their partners. Harries utilised both individual interviews and focus groups to explore the views and perspectives of parents who had their children removed into out-of-home care. The interviews were designed to be flexible. Interviews began as unstructured, oral story interviews and were then assessed based on the ease of conversation flow. If this interview technique was unsuitable for a particular interviewee, the interview would take on a semi-structured form, where the interviewer would prompt the interviewee to talk about specific aspects of their experiences and perspectives. Participants were also able to bring a support person to their interview. Three focus groups were also conducted, and included parents who had previously been interviewed separately, as well as new participants who had not previously featured in the research.

Hinton used face-to-face interviews with 20 families about their experiences of the child protection system. The interviews were generally conducted in the homes of the families, or another comfortable location, and the participants were allowed to bring a support person. A further six focus groups were conducted with a total of 27 parents. Following the research, support services were offered to the participants if required.

---

18 Harries, above n 5, 6, 9.
19 Ibid 5.
21 Ibid 9.
22 Hinton, above n 5, 141.
23 Ibid 142.
24 Ibid 141.
25 Ibid 142.
Taplin and Mattick’s study into the experience of mothers who had children in care while undergoing drug rehabilitation used structured interviews to explore their experiences of the child protection system, as well as demographic, social, mental health, and drug use information.\(^{26}\) The researchers also accessed the child protection records of the mothers.\(^{27}\)

Kiraly and Humphries explored the experiences of 20 mothers with children in kinship care through the use of semi-structured interviews.\(^{28}\) The researchers were aware of the power imbalance between the participants and the interviewer, and so took measures to reduce the effects of this imbalance. The interviews took place in comfortable environments, such as the participants’ homes, the participants were rewarded for their participation, and they received a copy of a progress report and the final report.\(^{29}\)

Other studies have investigated the perspectives of people with children in care with particular disadvantages. McConnell and colleagues implemented mixed methods to investigate the experiences of parents with mental health issues in the child protection system.\(^{30}\) The researchers reviewed court files and observed parents in court, although parents were not directly included in the study. Focus groups and interviews were also conducted with magistrates, legal representatives and child protection workers.

Qualitative unstructured and semi-structured interviews allow rich, detailed perspectives to be documented, and allow the interviewee to discuss the matters they feel are most important to their experiences.\(^{31}\) This contrasts with quantitative methods which investigate issues predetermined by the researcher to be important. Such research is typically more generalisable than qualitative methods but misses opportunities to understand important factors related to people’s unique experiences and circumstances.

Quantitative child protection data can give some picture of families’ and children’s circumstances but should be considered in unison with other qualitative evidence because of

\(^{26}\) Stephanie Taplin and Richard P Mattick, ‘Supervised Contact Visits: Results from a Study of Women in Drug Treatment with Children in Care’ (2014) \textit{39 Children and Youth Services Review} 65-72, 67.

\(^{27}\) Ibid.

\(^{28}\) Kiraly and Humphreys, above n 7, 108.

\(^{29}\) Kiraly and Humphreys, above n 7, 108.


particular methodological limitations. For example, certain families such as single parent families come to the attention of child protection authorities more than others. Policy and social factors influence data, for instance where there are different reporting and definitions across jurisdictions, cultural values and differing norms about child abuse. There may also be inconsistencies in how data is collected, for instance, in relation to methods and purposes of data collection and prospective versus retrospective collection.32

Evidence from previous studies
The handful of Australian and international studies of parents’ experiences have common findings: parents find court proceedings traumatic and alienating and feel unable to participate in these processes.33 Parents who have contact with child protection proceedings and processes experience guilt, fear, shame and a sense of powerlessness. They are unable to access adequate support and information during child protection proceedings and experience grief and trauma following children’s removal.

Australian research shows that parents’ lawyers are aware of their role in countering significant power imbalances between parents and the state in child protection proceedings.34 It is difficult to ensure procedural fairness in an adversarial system that has a high level of informality, such as no rules of evidence. Research has shown collaborative approaches between professionals and children’s parents ensure the best outcomes for children and families.35 In NSW, McConnell et al have researched the power imbalances for people with intellectual disability,36 and Ainsworth and Hansen have explored the difficulties parents have in managing the

complexities of child protection proceedings. These difficulties include finding and consulting with a solicitor, often a Legal Aid lawyer; assembling supporting documents in short time frames; problems associated with no application of the rules of evidence; challenges for parents complying with aspects of the adversarial system; parents’ lack of understanding or knowledge of any rules about what can and cannot be said during contact with children; and a lack of support services. In New South Wales, under section 93(3) of the Children and Young Persons (Care and Protection) Act 1998:

The Children’s Court is not bound by the rules of evidence unless, in relation to particular proceedings or particular parts of proceedings before it, the Children’s Court determines that the rules of evidence, or such of those rules as are specified by the Children’s Court, are to apply to those proceedings or parts.

Studies in other states support how difficult it is for parents when child protection departments provide incriminating evidence about parents that is not in accordance with the rules of evidence. In a study undertaken in Tasmania, lawyers and advocates commented on significant use of hearsay evidence by child protection services and the low level of proof required for something as fundamental as parents keeping their children. There are a few international studies of parents’ legal representation that demonstrate the complex requirements for the effective representation of parents. Guggenheim in the USA believes specialised training is necessary for lawyers who represent parents. In England and Wales, Shaw and colleagues


38 Hinton, above n 5, 82.


40 Ibid.
have raised the complex issue presented by recurrent removals of children from the same parents: this clearly demonstrate the need to understand how parents experience and respond to processes and proceedings.41

In summary, there is limited research into parents’ perspectives on legal processes or into how legal and administrative processes impact their lives and relationships with their children. Although both Australian and international studies have identified common parental responses to investigation, such as, feelings of powerlessness, fear and stigma, 42 parents voices about the impact of these systems are rarely heard by those in positions of power, and parents make only a marginal contribution to reforms to these systems. Previous research has recognised the lack of parents’ involvement in research, policy and practice and law reform in Australia.43 In prior studies, similar groups of parents have expressed appreciation of the rare opportunity to have their say in an area of deep personal interest to them.44 Research carried out with parents that asks them about their experiences, perspectives and ideas meets their needs to be involved, allows for public recognition of their collective loss, and can improve outcomes for children.

Evidence gap that prompted the research
Research into parents’ experiences can inform legal and professional practice with parents and help to generate a focus on outcomes for children who are the subject of legal proceedings and/or in care. It was particularly important to research experiences of parents in the Hunter region, which has one of the highest rates of children in out-of-home care in NSW, particularly when NSW has a higher rate than the national average. High rates, and variations across jurisdictions, are not well understood, but likely reflect the social and structural disadvantages described in this article. Reforms to the child protection system in NSW in 2014 increased the need for research into parents’ experiences as they participated to a limited extent in consultations about these changes.

Guggenheim, (2007, 2016)
42 See above n 5, n 6 and n 7
44 Kiraly and Humphreys, ‘A Tangled Web: Parental Contact with Children in Kinship Care’ above n 43, 108.
Research with parents in the Hunter Valley in 2016

Parents’ experiences of participation in child protection proceedings and processes has not been the subject of research in the past in the Hunter region. The Hunter region in NSW has one of the highest rates of children in out-of-home care in Australia, at a rate of about 1.8%, compared to a state-wide rate of approximately 1% and a national rate of 0.8%. Rates of children in out-of-home care are generally much higher in rural and remote regions.\(^{46}\)

The Context for the research: child protection reforms in New South Wales

NSW and other Australian child protection systems have traditionally been risk focused, with less emphasis on prevention, early intervention and support.\(^{47}\) In 2008 the Wood Special Inquiry into Child Protection Services in NSW instigated major changes to the system, recommending a move to a public health model and a staged transition of out-of-home care services to the non-government sector.\(^{48}\) This led to large numbers of non-government organisations providing casework services and out-of-home care, with increased funding for government child protection services and secondary services. Expenditure on out-of-home care services remained high due to higher numbers of children and young people remaining in in out-of-home care for longer. The spending on secondary services failed to impact these growing numbers and, in fact, may have contributed to the increases.

Major legal reforms commenced in NSW in October 2014, with permanency of care the key focus: although a consultation was undertaken prior to the changes, there was minimal input by parents. These reforms prioritised guardianship and adoption over long term care of children in out-of-home care. Both these options have consequences for parents’ ongoing relationships with children. These reforms were coupled with reforms intended to enhance caseworkers’ focus on parents, for instance, through the introduction of parent capacity contracts.

As discussed above, previous research in Australia and other countries has found that parents are systematically disadvantaged by processes in these systems, with very poor outcomes for parents and their ongoing relationships with children. There have been no similar studies in

\(^{45}\) Department of Family and Community Services, Government of New South Wales, above n 2.


NSW since the inception of the reforms. The research in the Hunter region was undertaken with a small group of parents in response to this gap, to better understand their experiences and concerns. The research was qualitative and adopted a family inclusion and children’s rights perspective in its design.49

**Family inclusion and a children’s rights perspective**

‘Family inclusion’ and ‘family inclusive practice’ – sometimes called ‘family engagement’ in the literature – are developing terms and were a key focus in this research. The practice involves a collaborative process for the meaningful inclusion of parents in their children’s lives, in ways that extend beyond the usual, often limited, contact visits. It is fundamental to maintaining important family and social connections for children in out-of-home care, and to supporting parents to be better parents, regardless of whether children remain in care permanently. The concept of family inclusion is important as parents have tended to be largely invisible as subjects of concern in practice and in the literature, once their children have entered care.

A children's rights perspective emphasises the importance of collaborative and meaningful practice with parents and families of children who are the subject of legal proceedings and where children are placed in out-of-home care. Children have a right not only to be protected from harm, but to be supported and cared for by their parents, and parents have responsibilities to uphold those rights.50 Children and young people in care have expressed their need to have a sense of belonging and relationships with their family of origin and others.51 Research shows that maintaining stronger relationships between parents and children leads to better permanency and stability during care,52 better outcomes leaving care, including less loneliness

---

49 For a full report on this research, see Nicola Ross et al, ‘No voice, no opinion, nothing’: Parent Experiences when Children are Removed and Placed in Care’ (Research Report, The University of Newcastle, 2017).


and more practical support,\textsuperscript{53} improved outcomes in adulthood,\textsuperscript{54} higher rates of restoration\textsuperscript{55} and stronger relational permanency, contrasting with a narrow measure of legal permanency.\textsuperscript{56}

**Research Design**

The research aim was to explore parents’ experiences and perspectives of child protection proceedings and processes, including processes associated with children’s removal, contact and support service for parents. A further aim was to examine if a policy of family inclusion was reflected in parents’ accounts of their experiences.

To provide insights into parents’ lives and perspectives a qualitative study was adopted, using conversational methods of interviews and focus groups. Lived experience research can raise key people’s awareness of others’ experiences. Otherwise, they may never hear those stories, or may hear them in pressurised work situations such as court proceedings, assessment, and children’s removal.

Recruitment can be a very significant challenge in studies with this parent group due to stigma associated with their disadvantage. Most parents for this study were recruited via social media and parent-peers, and some responded to emails and flyers sent via non-government out-of-home care services and parenting programs. The difficulty of accessing this parent group can be minimised by working with organisations that are set up to support this group of parents.

The most successful avenue of recruitment via services was through family support services, particularly where parenting programs targeted this particular group of parents. This research required a balancing act between protecting participants from further trauma and harm during the research and acknowledging their right to be heard and consulted about matters affecting them. It was also necessary to alleviate parents’ concerns about the potential repercussions of talking about systems they found disempowering. An important innovative approach was the inclusion of two parent consultants – with personal experience of child removal and out-of-


\textsuperscript{54} Mendes et al, above n 53, 368; Judy Cashmore, Marina Paxman and Michelle Townsend, ‘The Educational Outcomes of Young People 4–5 Years After Leaving Care: An Australian Perspective’ (2007) 31(1) *Adoption and Fostering* 50, 56.

\textsuperscript{55} Elizabeth Fernandez and Jung-Sook Lee, ‘Accomplishing Family Reunification for Children in Care: An Australian Study’ (2013) 35(9) *Children and Youth Services Review* 1374. This article discusses some findings from a four-year prospective longitudinal study of reunification process and outcomes.

\textsuperscript{56} Gina M Samuels, ‘A Reason, a Season, or a Lifetime: Relational Permanence among Young Adults with Foster Care Backgrounds’ (University of Chicago, 2008), 14.
home care placement – who assisted the researchers with framing questions for participants, analysing de-identified data, and co-facilitated one to two focus groups each with a researcher.

Focus groups are commonly used in research involving disadvantaged groups. Participants may feel empowered and supported while participating in focus groups, and the relaxed structure may enable participants to control the discussion and incorporate into the discussion topics that they feel are the most important.\(^57\) However, focus groups may not always be appropriate. Some participants may not feel comfortable engaging in public speaking in front of an audience of other participants and researchers; and may feel the need to censor their language and opinions, especially in focus groups where many strangers are present.\(^58\) This can potentially be avoided by composing focus groups of participants who have existing social relationships. The research team included parent consultants as co-facilitators of focus groups to help overcome some of the discomfort that parents might experience when discussing their lives with researchers who may have different life experiences.

Most of the interviews with parents were carried out in their homes, with some taking place by phone due to logistics or child care needs. An early attempt to put aside a day on a weekend to interview parents, with lunch and refreshments provided, was not successful. As in other studies of this kind, parents appear to be more comfortable talking about their experiences in their own homes. The first part of the interview was taken up with the collection of some basic demographic information, which helped researchers to understand the structure of participants’ families and their narratives. It also allowed basic information about the participant group to be collated for the purposes of the research.

**Participants**

A total of 18 parents participated in interviews, 13 women and five men. Eight of those parents also participated in focus groups. The group included three couples, which meant participants represented 15 families. Their average age was 35 years, and eight of the parents were caring for children at home, either on their own or with their own parent. The three couples did not have children living at home. The participant families had a total of 50 children, 34 of whom


\(^{58}\) Deborah J Warr, “‘It Was Fun … But We Don’t Usually Talk About These Things’: Analyzing Sociable Interaction in Focus Groups” (2005) 11(2) *Qualitative Inquiry* 200, 202.
had been removed. Of those children, 27 were still in care (all kinship or foster care), six had been restored (four to the same family), and one had left care once over 18 years.

**Findings**

The study provided rich data on the experiences and perspectives of parents. Five primary themes emerged related to power and inclusion, professional relationships and relating, parent-child relationships and attachment, grief and loss, and identity. Legal experiences appeared across all of the themes, although predominantly in ‘power and inclusion’ and ‘professional relationships and relating’. Although identified separately, power and inclusion ran through all themes as a major feature or explanation of most parents’ personal and system experiences. The following describes key findings related to legal proceedings, processes and practice.

**Not a level playing field, even with a lawyer**

Most (but not all) parents in the study commented positively about the assistance they received from their lawyer. The lawyer’s ability to communicate respectfully with them and put them at their ease was very important. Honest and clear feedback was appreciated, so that parents knew where they stood.

> My lawyer was brilliant, she was brilliant. She was one of those ones that was quite honest with me about what was going to happen and how it was going to go.

Where parents expressed dissatisfaction, this was where they felt judged or neglected by their lawyer, or felt the lawyer lacked the necessary expertise in this specialised area.

> …the first one didn't listen, wouldn't get affidavits in on time, wouldn't meet with me to discuss court. (My new solicitor) just listens. I think that's the biggest part in solicitor-client representation. If the solicitor doesn't listen to you then they don't know what you really want. They're going off their own back, agreeing to almost everything and that's not what I want.

One young parent had a very poor experience with her lawyer. He was judgmental and she struggled to understand him; he did not present her case well. This had serious consequences – her baby was removed. Although she tried to access resources she needed, to facilitate the care of her child, she was unable to do so when her the child was removed and was unable to understand why this occurred.
[It] confused me because I did the best I could and it still wasn’t enough for them…I had a house, was waiting for the psychologist. I engaged with the young parent social worker.

Later, when she was provided with the services she had asked for (primarily, counselling from a psychologist), she was able to successfully parent the next child she gave birth to.

Parents said that legal representation by itself did not ensure that there was a level playing field. They felt the legal processes worked against them, for instance by allowing the lawyers for Family and Community Services to tell lies or partial lies about them, without an opportunity to question these lies. Parents’ lawyers also had to comply with the legal processes: according to some of the lawyers, these worked against parents who “didn’t have a lot of rights”.

Even the lawyer said you don’t have a lot of rights when you’re a parent and your child’s been removed.

Some of the reasons for this relate to court processes, such as the non-application of the rules of evidence, discussed above.

**Negative characterisation of parents; isolation in the court process**

Parents found court to be an intimidating, frightening and humiliating experience. They felt the process was not only impersonal, it was dehumanising. The judge didn’t look at them, but only dealt with the lawyers.

…they don’t even really acknowledge that you're in the room. They'll ask and the judge will say is the mother present but she doesn't look at you...

The process left out vital parts of their experience – for instance where domestic violence issues were not properly acknowledged.

Court was awful. It was belittling, worst feeling ever. The judge, looking at you like you're the biggest loser, let your children down…I had nowhere else to go.

My ex, we had a domestic violence relationship, so it made me feel scared as well when I was in the courthouse with them all and them all lying. It was horrible…the anxiety and everything you get before you walk into that court house is horrible.

One parent discussed having to breastfeed a newborn in court; no allowance was made for her to do so and she found the experience of the judge “seeing her boobs” humiliating.
Parents were not allowed to engage in an active way with proceedings and often felt marginal and irrelevant, even though the outcomes were vitally important to them. Many expressed the concerns they felt when negative characterisations were made about them in court.

I didn't feel really part of the process. I was being talked about but I wasn't being spoken directly to. It was a judge talking to DOCS and the lawyers... Some evidence from community services wasn't true or not completely true. It was hard to not stand up and say, that's not fair, that's not true.

Despite their lawyer’s presence, parents did not feel powerful enough to question what they saw as false evidence or lies, as no rules of evidence applied. They perceived that the more powerful position of FACS was used to dominate court processes and eventually to determine outcomes, without proper time and space to address issues such as what they had achieved in working towards providing adequate support for their children.

**Removal began a process that distanced parents from children**

Parents found removal of their children extremely traumatic. They expressed how difficult it was without support or information about the process, such as when they could see their children next, or what they had to do in order to have their children restored. They often had to attend court soon after removal and this meant they needed to find a good lawyer quickly – they often didn’t know where to start. With limited information and resources, they struggled to find ways to do what was necessary to have their children returned. Programs they were told they needed were often not available or had rules that excluded parents whose children were not living with them, so there was a sense of catch 22.

I hear it from a lot of parents that when their kids are taken into care, they don't know what to do, who to see, or parenting programs they can do…in the court process they say, you need to do X, Y, Z parenting courses…you've got to wait for them to come up. I think it would be good if when a child's taken into care…at least putting an information pack in their hands…going, this is what's next, this is the procedure.

Parents became aware of the importance of bonds between parents and children to the ultimate decision about whether children would be returned to their care. They were often advised by their lawyers not to challenge FACS early decisions to remove the children – and for many parents, this began a process of them being distanced from their children. Their relationships were severely impacted by limited contact of an hour once or twice a week, often in a foreign setting, with unfamiliar supervisors.
…they really try and break that bond between you and the kids. Then that’s one of their reasons for not letting you have the kids back because they’re settled where they are and you’ve only been seeing them once every fortnight.

In relation to one child who was reunified with his parent, contact was ordered by the court prior to final orders being made to return him to the parent’s care, and this included telephone calls. However, the carer failed to follow these orders and there was nothing the parent could do to enforce the orders – or as can be seen here, to keep up to date with their child’s medical situations.

The carer forgot some weeks totally. I would sit next to the phone and wait and wait; the phone call just never came. I wasn’t involved in any medical care until he was restored.

These situations caused children and parents further harm, which had to be sorted out following restoration.

**Discussion – Law, policy and practice**

This research adds to existing sources of child protection knowledge and, importantly, provides rich and powerful examples of the lived experience of parents who had children removed and placed in out-of-home care. The findings clearly echo what other researchers have found about the powerlessness and angst of parents in child protection processes and legal proceedings. As in similar studies, parents in our study overwhelmingly said they did not receive help when they sought it, did not understand how to participate in child protection processes and proceedings, and did not believe they were on a level playing field when the court was deciding their children’s future. They found it difficult to maintain suitable contact with their children while they were in out-of-home care, pending a final order being made about whether children were to return home permanently, which significantly affected their relationships with their children.

The research team published a report and followed up the report launch with publicity and involvement in academic and policy forums to promote the findings and to change how these parents are viewed.\(^\text{59}\) The team is also communicating with key stakeholders about how parents

---

\(^{59}\) The research team have submitted articles to a number of journals and clearinghouses and have made submissions in relation to proposed legislative reforms in 2017. Presentations were made in 2016 to the Australian Institute of Family Studies Bi-annual Conference and to the Association of Children’s Welfare Agencies Annual Conference, and in 2017 to the Second International Conference on Non-Adversarial Justice, Sydney and at Oxford University (UK).
experience legal and other processes, highlighting how critical parents are in their children’s lives, no matter where children reside. The research team are particularly interested in how practitioners will apply the findings, and have used practice forums, conferences, surveys and workshops to communicate with practitioners about how they see the findings and what they might do with them.60

This group of parents have key insights which need to be listened to if reunification of children with their families is to be accorded the priority it is given in the legislation. Evidence from this research can help change negative community attitudes and stereotypes about these parents. It can raise awareness of the significance of social and economic disadvantage to children’s safety and wellbeing, as well as parents’ capacity to care for their children and to engage in legal proceedings. It can inform the introduction or expansion of alternative pathways and more collaborative approaches to work with parents that focus on both heightening parents’ participation and strengthening children’s identity through minimising disruption to their connections with families of origin.

The findings of this research can also inform related legislative and procedural changes. It raises important questions about the level of support for parents once children are removed from families. These questions include; should children be removed if relevant services have not been provided? What role does and could the court play in ensuring that children are not removed before support services have been offered? What is needed to support more collaborative processes between parents, child protection workers, staff from out-of-home care agencies and carers? Asking these questions can help lead organisations to develop processes that acknowledge the contexts of children’s and families’ lives across the spectrum of child protection interventions, from prevention or early intervention to legal proceedings and participation in court. This can ensure better-informed and evidence-based decision making about whether or not parents and their children can remain together.

Research such as this increases the visibility of parents and their lived experiences and can recalibrate how legal systems and professionals think about parents. This can humanise parents and normalise their responses to intervention, legal practitioners, and helping professionals. By extension, this allows parents and their perspectives to be included in practice and policy

60 To date, the research team have presented findings at a Hunter-based out-of-home care forum in March 2017, to a Family and Community Services caseworker conference in Sydney in September 2017, an Association of Children’s Welfare Agencies restoration forum in Sydney in July 2017, and as part of an Australian Institute of Family Studies webinar in August 2017.
development, to support more effective consideration of questions about the law-in-action and forming and sustaining quality parent-practitioner relationships.

**Conclusion**

Due to the stigma parents experience, and because the need for research must be balanced with respect for parents’ disadvantages, this group of parents is not easily accessed for research purposes. The use of a multidisciplinary team made up of academic and practising lawyers and social workers and parent consultants was an important feature of the qualitative research design and helped us to reach out to parents in a way that appeared to meet their need to talk about their experiences. It also modelled what collaborative practice between mixed disciplines, and parents and professionals, could look like in practice.

This research is one example of qualitative research which investigates the impact of the law on a disadvantaged population. We examined how those disadvantages impact on parents’ ability to interact with and participate in child protection processes and proceedings. Social justice for children and their parents requires that policy makers and practitioners respond to the needs of this group by better supporting their participation in processes and court proceedings. They can do so only if better research evidence is available to them about how this group of parents experience their interactions with processes and proceedings, to provide a basis for further reform and improvement. Finally, and most importantly, this research can enhance opportunities for restoring children to parents where appropriate, which is the first goal in all child care proceedings, including those in New South Wales.
Methods to Evaluate Justice Practices in Eliciting Evidence from Complainants of Child Sexual Abuse

Jane Goodman-Delahunty*, Eunro Lee**, Martine B Powell*** and Nina Westera****

Abstract

To facilitate evidence-based practice, following implementation of a reform, evaluation of field practice involves rigorous and scientific methods. The present article reviews the methods and implications of 17 studies commissioned by the Royal Commission into Institutional Responses to Child Sexual Abuse to evaluate how alternate measures are being used and how complainants are being questioned about child sexual abuse in Australian criminal justice proceedings. This evaluation used qualitative and quantitative methods to assess the processes and practices applied to manage the witnesses’ psychological distress and vulnerability to elicit more reliable and credible evidence. The topics and areas of interest were stakeholder perceptions and views of the use of alternate measures, factors considered in the processes, and challenges in police interviews and courtroom questioning. Data sources included video recordings of interviews by police and CCTV cross-examinations and trial transcripts supplemented by stakeholder interviews, a survey, and an online experiment.

In particular, the evaluation findings highlighted a range of practices based on unsupported assumptions about victim memory and behaviour, judicial instructions to child complainants, cross-examination strategies, judicial interventions, and shortcomings in the quality of video recordings of police pre-interview and CCTV cross-examination. In this article, following a review of the studies, implications of the findings are discussed to inform evidence-based practice and research in eliciting evidence from complainants of child sexual abuse.

Key Words: evidence-based practice, child sexual abuse, elicitation of evidence

* Research Professor, Faculty of Business, Justice and Behavioural Sciences, Charles Sturt University. Author contact: jdelahunty@csu.edu.au.
** PhD, Lecturer, School of Health & Biomedical Sciences, College of Science, Engineering & Health, RMIT University. Author contact: eunro.lee@rmit.edu.au.
*** Professor, Centre for Investigative Interviewing, Griffith University. Author contact: martine.powell@griffith.edu.au.
**** Research Fellow. Griffith University, Australia.
**Introduction**

Delivery of justice is critically dependent upon the quality of trial evidence, for which a range of legislative and judicial procedures exists to facilitate the elicitation of evidence of the best quality. This is particularly salient in the criminal justice sector regarding the evidence of the most vulnerable of all witnesses, namely complainants of child sexual abuse. Child sexual abuse is difficult to prosecute and has one of the highest attrition rates of all criminal offences.\(^1\) Although exact figures are difficult to come by, estimates suggest that only about eight to nine per cent of child sexual assault cases reported to police are prosecuted.\(^2\) Part of the difficulty in prosecuting these cases is that the offending is often hidden from public view, leaving only the complainant’s evidence to establish the defendant’s guilt beyond reasonable doubt.\(^3\)

A substantial body of research has established that complainants of child sexual abuse experience severe anxiety, distress, and psychological difficulties during legal processes, due to the witnesses’ developmental characteristics, the impact of confronting the offender and adversarial cross-examination in court.\(^4\) These findings co-occurred with low reporting rates to police (8-9%), high attrition rates during the investigative process (81-85%), low prosecution rates (15%), and low conviction rates at trial (8%).\(^5\) These trends have emerged in multiple jurisdictions nationally and internationally.\(^6\) As a consequence, miscarriages of justice may occur at various stages of the criminal justice process.

---


Figure 1. Phases of the Criminal Justice Process

In the Australian legal system children were traditionally viewed as an unreliable class of witness. Because of their age, they were presumed incompetent by law, and their evidence had to be independently corroborated. For child complainants, giving evidence at trial about sexual abuse was a potentially traumatic process, resulting in great anxiety and stress. In addition, prosecution outcomes were poor, leading many complainants (or those responsible for their welfare) to opt out of pursuing a criminal prosecution. These factors led to concerns that the process for giving evidence in criminal trials was working against the interests of justice for children, and that prosecutions failed because the court process was ill-suited to children. In recognition of this fact, a range of special measures, also known as ‘alternate measures’, was
introduced by courts in Australian and other jurisdictions to assist complainants in presenting their evidence in legal settings.\(^\text{13}\)

The experiences and satisfaction of vulnerable and intimidated witnesses were systematically assessed before\(^\text{14}\) and after\(^\text{15}\) special measures were implemented in the United Kingdom in 1999. Data were gathered by means of a survey administered in both 62 Crown Courts\(^\text{16}\) and 48 Magistrates’ Courts\(^\text{17}\) for this purpose.\(^\text{18}\) Some survey participants (30%) also took part in interviews. The researchers found that the overall level of satisfaction by the cohort of vulnerable witnesses was significantly lower than that of other witnesses in the criminal justice system. The findings of this UK study suggested that the effectiveness of alternate measures implemented in Australia could be assessed using a similar suite of research methods.

In 2013, the Royal Commission into Institutional Responses to Child Sexual Abuse (hereafter ‘the Royal Commission’)\(^\text{19}\) engaged a team of law and psychology researchers to investigate how these special measures were being utilised in Australian jurisdictions.\(^\text{20}\) This Royal Commission project further explored whether improvements were needed to enhance the quality, reliability and credibility of evidence from complainants of child sexual abuse.\(^\text{21}\) The present article reviews the methods employed in the Royal Commission project comprising a multi-study research evaluation, with regard to a range of Australian special measures available in cases of child sexual abuse.

In the Royal Commission project, a variety of qualitative and quantitative methods were applied in a total of separate 17 studies to conduct a robust evaluation of the effectiveness of these alternate measures in eliciting evidence from complainants and vulnerable witnesses. Assessment of apparent benefits and effects of alternate measures revealed by empirical legal

\[^{13}\] Pigot, above n 8.
\[^{14}\] \(N = 552\).
\[^{15}\] \(N = 569\).
\[^{16}\] \(N = 62\); 86 for selected phases.
\[^{17}\] \(N = 48\); 94 for selected phases.
\[^{18}\] Hamlyn et al, above n 9.
\[^{21}\] Ibid.
studies is contingent upon the research methods employed. The present article provides an overview of multiple methods employed in the 17 studies conducted in the course of the multi-study research in the Royal Commission project. The goal of the research was to determine whether the processes and practices implemented were successful in managing the witnesses’ psychological distress and vulnerability, so that more reliable and credible evidence is elicited.

Drawing on social science and psychology research approaches, the research team collected several forms of data from diverse sources. Notable sources were video recordings and transcripts of actual pre-recorded complainant interviews by police and of CCTV cross-examinations, plus transcripts of real child sexual abuse trials. These data provided unique insights into actual practices in three primary nominated jurisdictions, namely NSW, Victoria, and WA. In some studies, additional data from Queensland and Tasmania were included. In addition, the research team gathered supplementary data via stakeholder surveys, interviews and focus groups with criminal justice professionals including judges, prosecutors, criminal defence lawyers, police officers, and witness support staff. Complementary methodologies used for data analysis were derived from qualitative thematic analysis of interviews, content analysis of trial transcripts as well as quantitative statistics using analysis of variance, χ²-tests, factor analysis, and mixed model analyses of survey and experimental data. The foregoing methods comprised a rich and comprehensive evaluation of practice.

The multiple methods used in the Royal Commission project are illustrative models of evidence-based practice that can be applied in other legal sectors and to other legal issues as well as to complaints of child sexual abuse. One of the main advantages of including several different methodologies is that the findings converge to enhance our understanding of the matter. As a result, any limitations associated with one methodology are mitigated by the commonality of findings across methodologies. In this article, an overview of the methods used is outlined.

Before the specific research methods are discussed, policy changes permitting special measures for complainants of child sexual abuse are reviewed. In the conclusion, future applications of research methods in law and justice practice are discussed.

---

Special Measures for Child Sexual Abuse Complainants: Research and Practice

Starting in the 1980s, in Australia and other overseas jurisdictions across the globe, special measures were implemented to take the evidence of child complainants, such as the provision of support persons, and cross-examination via CCTV. Whether these measures, individually or collectively, accomplished the intended aims of reducing anxiety, intimidation, and frustration incurred by standard court procedures was untested. Accordingly, the Royal Commission nominated three primary jurisdictions to investigate the effectiveness of alternate measures, namely New South Wales (NSW), Victoria (Vic) and Western Australia (WA).

The discrete phases of the criminal justice process depicted in Figure 1 were examined in a series of different studies incorporated in the Royal Commission project. For instance, Study 4 analysed case files maintained by prosecutors to report on documented considerations of special measures in the period after a case was referred for prosecution, whereas Study 5 analysed the use of special measures at trial during the complainants’ evidence in-chief, cross-examination, and re-examination.

A combination of qualitative and quantitative methods was used in the programmatic evaluation research. Data sources included prosecutors’ files, video recordings of police interviews and CCTV cross-examinations, and trial transcripts. Each data source, the type of research method, and the topics of the respective studies addressed by the evaluation are summarised in Table 1.

Table 1. Data Sources, Qualitative and Quantitative Methods Applied to Evaluate Special Measures for Complainants of Child Sexual Abuse.

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Type and Topic of Studies</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative studies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

23 Mandy Burton, Roger Evans and Andrew Sanders, ‘Are Special Measures for Vulnerable and Intimidated Witnesses Working? Evidence from the Criminal Justice Agencies’ (Online Report No 01/06, Home Office (United Kingdom), 2006); Judy Cashmore, above n 10; Judy Cashmore and Marion Horsky, ‘The Prosecution of Child Sexual Assault’ (1988) 21(4) Australian & New Zealand Journal of Criminology 241; Emma Davies and Kirsten Hanna, Pre-Recording Testimony in New Zealand: Lawyers’ and Victim Advisors’ Experiences in Nine Cases’ (2013) 46(2) Australian & New Zealand Journal of Criminology 289; Robert A Nash et al, ‘Remembering Remotely; Would Video-Mediation Impair Witness’ Memory Reports?’ (2014) 20(8) Psychology, Crime & Law 756. For instance, in Victoria the Criminal Procedure Act 2009 (Vic) (‘CPA’) s 370 was amended in 2012 to prohibit the accused from attending the room in which the complainant’s evidence is taken (see CPA s372(1)(a) and (b)).
METHODS TO EVALUATE JUSTICE PRACTICES

<table>
<thead>
<tr>
<th>Interviews</th>
<th>Professionals’ views on how to improve evidence-taking</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case files</td>
<td>Prosecution case file review</td>
<td>4</td>
</tr>
<tr>
<td>Police interview transcripts</td>
<td>Police interview practices</td>
<td>8</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>The labelling of recurring abusive events</td>
<td>10</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>The utility of alternate measures</td>
<td>5</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Courtroom discussions about police interviews</td>
<td>9</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>The labelling of recurring abusive events</td>
<td>10</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Judicial directions to child complainants</td>
<td>11</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Assessing truth-lie competency</td>
<td>12</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Courtroom questioning</td>
<td>13</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Non-supported assumptions about victim memory and behaviour</td>
<td>14</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Cross-examination strategies observed</td>
<td>15</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Questioning about and handling of inconsistencies</td>
<td>16</td>
</tr>
<tr>
<td>Trial transcripts</td>
<td>Judicial intervention</td>
<td>17</td>
</tr>
<tr>
<td>Minutes of the NSW Sexual Assault Review Committee</td>
<td>Content analysis regarding alternate measures</td>
<td>6</td>
</tr>
<tr>
<td>Video recordings</td>
<td>Analysis of the quality of pre-recorded police interviews, CCTV cross-examinations and non-verbal features of the speakers</td>
<td>7</td>
</tr>
<tr>
<td><strong>Quantitative studies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey</td>
<td>Professionals’ experiences with and views of special measures</td>
<td>2</td>
</tr>
<tr>
<td>Experiment</td>
<td>Factors that influence perceptions of the quality of cross-examination</td>
<td>3</td>
</tr>
</tbody>
</table>

A total of fifteen qualitative studies and two quantitative studies were included in the evaluation program conducted for the Royal Commission. These are outlined below.

**Qualitative Study 1: Interviews of criminal justice professionals.**

To examine how well evidence is taken from complainants of child sexual abuse, the most conventional research methodology in social science, an interview study, was conducted with
criminal justice professionals most closely acquainted with child sexual abuse complaints. This approach was selected to achieve a detailed and in-depth understanding of any barriers and limitations surrounding the use of special measures in practice.\textsuperscript{24} By exploring stakeholders’ individual experiences and perceptions in a narrative format, interview data can provide an insights and a ‘closer’ investigation of the topic than is achieved by means of quantitative information such as frequencies and proportions of uses of alternate measures in each jurisdiction, or statistical reports of conviction rates or survey results.

To conduct the interview study, the research team contacted police departments, public prosecution offices, law firms, barristers’ associations, courts, and witness assistance agencies in the target states. Judges, prosecutors, lawyers, and witness support staff were invited to participate in interviews to explore their day-to-day experience, regarding special measures for complainants of child sexual abuse. A semi-structured interview schedule was sent to each participant prior to the interview. Thus, a pre-determined set of interview topics was explored with all participants.\textsuperscript{25} Furthermore, using broad open-ended questions, these topics probed for information about preparing a complainant for trial, uses of special measures to question the complaint and present evidence. A total of 43 criminal justice professionals participated in the study, and were interviewed in person or by telephone. The interviews focused on how well special measures are being used in their everyday practice with child and vulnerable adult complainants of childhood sexual abuse. The participants’ thoughts and suggestions to improve criminal procedures and alternate measures were further explored.

All interviews were audio-recorded and transcribed for analysis using an inductive, ‘bottom-up’ approach,\textsuperscript{26} rather than a pre-existing theory. This approach affords the opportunity to include unexpected themes and agendas that researchers may not have anticipated, theorised or hypothesised beforehand.

**Qualitative Studies 8 and 10: Analyses of police interview transcripts.**

Detailed factual information on how complainants of child sexual abuse are interviewed when they first encounter representatives of the criminal justice system are difficult to obtain as the


\textsuperscript{26} Tom Wengraf, *Qualitative Research Interviewing: Biographic Narrative and Semi-Structured Methods* (Sage Publications, 2001).
records of interview are typically confidential. In this evaluation, police interview practices were evaluated in Study 8 by analysing a sample of 118 transcripts of police interviews gathered from the three nominated states. This method allowed the researchers to identify gaps between the recommended interview guidance and actual practice with alternate measures. Through content analyses of the police interview transcripts, researchers discerned the proportions of open-ended questions, leading questions, and non-verbal aids used in practice. One finding was that recommendations for rapport building using open-ended questioning were not adhered to by a majority of the police interviewers. Thus a training need emerged. Likewise, in Study 10, the labelling of recurrent abusive incidents by the same offender was specifically analysed in police interviews of 23 complainants. When responding to questions about recurring abuse incidents over a span of time, confusion can arise when police, lawyers and child witnesses use different terminology and labels to refer to the same and similar types of abusive acts on different occasion. Thus, sensitivity to terminology and labels that are developmentally appropriate for and used by children can be critical in eliciting accurate evidence throughout the criminal justice processes from the initial police report to cross-examination. More effective and useful labelling of each abuse incident will assist police in eliciting more reliable and credible evidence from complainants. This analysis highlighted some limitations in the practice, in terms of the source of the labels used and their consistency across interviews of the same complainant. A beneficial outcome of this evaluation method is the potential to guide the future training focus and professional development of police interviewers on rapport building and on labelling of recurrent abusive events to close the observed gaps between evidence-based recommendations and practice.

**Qualitative Study 4: Prosecution case file review.** To further evaluate the impact of the witness procedural reforms in Australian jurisdictions, the effectiveness of the reforms for complainants of child sexual abuse was examined by assessing how the measures were considered during the prosecution phase. This was accomplished by means of a manual review of prosecutors’ files in a sample of 60 most recent child sexual abuse cases lodged after 2010 in NSW, Victoria, and WA. A coding protocol was developed by the project researchers to analyse these considerations in light of case characteristics. The coding generated quantifiable descriptors to compare the frequencies and proportions of the uses of special measures such as CCTV, cross-examination, or assistance by an intermediary. Comparisons were also conducted of contemporary versus historical claims and institutional versus non-institutional claims.
These quantitative analyses were supplemented by further in-depth qualitative methods using thematic analysis. The analysis classified text excerpts from the prosecution file entries, for all comments recorded that showed explicit considerations surrounding the use of the special measures. The purpose of the systematic thematic analysis was to identify the reasoning and motivation for the use or non-use of eligible alternate measures. Examples of prominent themes that emerged were the complainant’s needs, legislative compliance, and logistics pertinent to the special measures. Content analysis coding permitted the frequencies of each theme to be analysed by complainant type, age, and case type. These results of the content analysis provided indirect measures of the importance of each theme to the legal practitioners in the course of litigating these claims. The results shed light on the complainants’ preferred methods for giving evidence and tensions between the prosecution and the complainant’s caregivers. In summary, the file review method using a coding protocol and systematic thematic analysis provided unique empirical evidence that will be useful for policy reform evaluation.

Qualitative Study 6: Analysis of the Minutes of the NSW Sexual Assault Review Committee.

Broader and wider perceptions of the use of alternate measures for complainants of sexual abuse held by various agencies and stakeholders were explored in Study 6 by examining archival records of the NSW Sexual Assault Review Committee (SARC), founded in 1993. This Committee is comprised of representatives from various government and non-government agencies in the justice, health, and community sectors who deal with sexual assault claims. The committee has met quarterly and regularly discussed NSW legislative reforms on evidential procedures in child sexual abuse cases. The minutes of all SARC meetings held in the period 1993 to 2014 were reviewed and content pertaining to alternate measures was analysed using an inductive qualitative method.

This desk review revealed multiple themed categories documenting various technical, administrative and practical logistical barriers encountered in implementing alternate measures. By analysing the minutes across the period spanning 20 years, chronological perspectives and insights were gained and staff training needs were highlighted. The findings provided a richer context on the progression of the uses of alternate measures in practice in a key Australian jurisdiction, compared to other studies using data from sources within the justice sector.

Qualitative Studies 5, 9-14 and 17: Analyses of trial transcripts.
Trial transcripts provide invaluable objective information on how the legislative reforms are implemented on the use of alternate measures for complainants of child sexual abuse. This method of trial transcript analysis contrasts with and supplements the subjective nature of interview responses. Although trial transcripts are a repository of objective information about a targeted practice, these data are typically difficult to access due to the sensitive nature of child sexual assault trials, the closure of many of these court proceedings to the public to protect the confidentiality and anonymity of the child complainants, and the costs of transcription.

The Royal Commission facilitated research access to a sample of 156 trial transcripts from the three target jurisdictions. Using these data, Study 5 evaluated which alternate measures were used, their prevalence and identified challenges in the nature of their use. Utilising a coding scheme, the resulting analyses revealed the proportion of complainants for whom specific alternate measures were used, by age group and jurisdiction. Uses of all available alternate measures were analysed in this manner, including instances of the removal of judges’ and lawyers’ gowns. Challenges and problems arising when using alternate measures were specified through these analyses. For example, the nature and frequencies of technological issues encountered in recording, editing, and playing of pre-recorded interviews in lieu of in-person evidence in chief by complainants were analysed by jurisdiction. Similarly, common problems arising with the support persons were identified and summarised.

A series of additional studies about alternate measures were conducted based on the trial transcripts. For example, Study 9 examined in-court discussions about issues pertaining to evidence in the form of pre-recorded police interviews, based on a sample of 96 trial transcripts. In these analyses, open coding was used to capture the content of the discussions line by line. Then, common topics were categorised, and the discussion content was coded by category. The results illuminated facets of the police interview structure and the procedure, as well as legal and technical problems in some interviews.

Using a similar trial script analysis method, judges’ directions to child complainants were analysed in a sample of 52 randomly selected trial transcripts from the three nominated jurisdictions. The researchers coded 11 distinct categories of judicial instructions. The major topics were ground rule instructions, the scope and length of the directions on the ground rules, and question formats, truth-lie competency, courtroom questioning, non-supported

assumptions about victim memory and behaviours, and judicial interventions during the trials. For instance, ground rule instructions at times included guidance on the “rule of don’t know” such as “I might ask a question and you don’t know the answer. Just say ‘I don’t know’.

A significant evidentiary concern in the justice sector is the use of traditional adversarial cross-examination strategies and tactics used by criminal defence lawyers. Prior research showed these strategies were rated as unfair and subsequently detrimental to the quality of the evidence elicited from complainants of child sexual abuse.28 By applying a coding scheme developed in previous studies,29 each line of cross-examination in the trial transcripts was coded and analysed in Study 15. The strategies and tactics used challenged the reliability of the complainant’s memory or the plausibility of the reported offence where the case facts included delayed reporting or the absence of resistance by a complainant at the time of offence. Despite empirical research showing that the assumptions implicit in the lawyers’ tactics are not evidence-based, the rates of the deleterious strategy use in the study sample ranged mostly above 90%, across jurisdictions, encompassing both children and adolescents. A strength of this research method is the objective evidence it yielded of actual courtroom practices.

A related and no less significant issue is the emphasis at trial on inconsistencies in the evidence elicited from complainants of child sexual abuse. This topic is central to many cross-examinations. A thorough and extensive review of memory research and developmental psychology demonstrated that minor inconsistencies are features of all human memory and rarely negate the reliability of the gist of reported abusive events.30 Drawing on this research evidence, Study 16 analysed details of the types of inconsistencies raised in cross-examining complainants of child sexual abuse in the three target jurisdictions. The categories included the content of common inconsistencies such as reports of when the abuse occurred, as well as


29 Sarah Zydervelt et al, ‘Lawyers’ Strategies for Cross-Examining Rape Complainants: Have We Moved Beyond the 1950s?’ (2016) 57(3) British Journal of Criminology 551.

sources of the inconsistencies. Based on these qualitative features, quantitative statistical analyses were conducted. The analyses included Chi-squared tests on the frequency with which these issues were raised with complainants in different age groups as well as analysis of variance and t-tests on the mean frequency of strategies and tactics used in these groups. This study comprises an exemplar of multiple methods, integrating qualitative and quantitative analyses to evaluate common but controversial practices in the criminal justice process.

**Qualitative Study 7: Video analysis of pre-recorded police interviews and CCTV cross-examinations.**

A number of special measures for complainants of child sexual abuse entail the presentation to juries of remote evidence presented visually in the form of video recordings of police interviews and CCTV evidence. The quality of the recordings and CCTV video displays has been a topic of concern and a potentially detrimental element in the adoption of these reforms. In Study 7 of the research evaluation project, a total of 102 electronic pre-recorded police interview videotapes and CCTV cross-examinations were systematically rated by trained coders. The resulting analyses revealed, for instance, that the audio and image resolution quality was poor or substandard in 17% of the cases. These results indicated that additional resources and effort should be focused on resolving this issue to improve the use of alternate measures for complainants of child sexual abuse. Additional coding was conducted of non-verbal features of the speakers (e.g., emotional expressions, eye contact, whether the speaker was mostly silent or actively interacted). A strength of this research method is the detailed objective information it provided to guide the allocation of further resources and policy development to advance the use of alternate measures in court by standardising the protocols for video recordings and remote CCTV transmission.

**Quantitative Study 2: A survey of criminal justice professionals’ views and perceptions of alternate measures.**

As a means of evaluating how a system is functioning, quantitative methods such as surveys of the practice professionals most familiar with the system can increase the robustness of the evaluation by canvassing the views and perspectives of a larger number of participants.

---

Compared to the in-depth and detailed subjective data gathered in interviews, an online survey described in Study 2, compared the uses of alternate measures implemented in three target jurisdictions. A total of 335 criminal justice professionals were recruited in five distinct groups to participate in the survey: judges, prosecutors, defence lawyers, police officers, and witness support staff. Results showed that their experiences with and training for child sexual abuse cases ranged broadly.

Statistical analyses were conducted on participants’ ratings of the reasons for non-use of alternate measures (e.g., logistics, credibility concerns) as well as ratings of the perceived impact of expert witness testimony, and the professionals’ views on best procedures for complainants of child sexual abuse. Collectively, these ratings provided insight into current practice. Qualitative analysis conducted on responses to open-ended questions using the software NVivo were useful in grouping opinions on expert witness testimony and suggestions for professional training on child sexual abuse. By including open-ended questions in the online survey, qualitative analysis complemented the statistical outcomes achieved.

The survey method delivers a more generalised understanding of the views and practice perceptions from a larger number of professionals than can be obtained from interviews or focus groups. The survey results can also be compared to the interview results. These two methodologies are distinct in the depth and breadth of the responses obtained regarding key issues. A strength of the survey method is that a more comprehensive perspective is derived of legal practices from a larger number of professionals, permitting new insights in evaluating the implementation of the alternate measures. A critical methodological point in survey administration is the nature of the sampling method. To increase the generalisability of the findings, random sampling or stratified sampling is required, so that potentially confounding variables (e.g., participant gender or professional role) are numerically balanced. A rigorous sampling method increases the representativeness of the results and minimises errors in the inferential statistics.\textsuperscript{32} Measures of participants’ views and perceptions are typically validated to ensure rigorous reliability and validity of the survey outcomes.\textsuperscript{33} In study 2, convenience sampling rather than random sampling was feasible to engage practising criminal justice professionals with a heavy work load.

\textsuperscript{32}Carl-Erik Särndal, Bengt Swensson and Jan Wretman, \textit{Model Assisted Survey Sampling} (Springer, 2003).
Quantitative Study 3: An online experiment on the factors influencing the fairness and appropriateness of cross-examination of complainants.

One of the most powerful empirical and scientific methods, an experimental design, was applied in Study 3. This online experiment involved the same 335 justice practitioners who participated in Study 2 to explore factors that influence perceptions of the quality of cross-examination of a complainant. A mixed model was used for this study. Specifically, each participant read two vignettes, in which three legal factors of interest were systematically varied as between-subject variables. Participants were randomly assigned to one of 8 different vignette conditions, in a 2 (victim age 10 vs 16 years: contemporary vs. historical) x 2 (questioning type: appropriate vs. inappropriate) x 2 (judicial intervention vs. no intervention) design. The complainant’s gender (male vs. female) and the two cross-examination vignettes were counter-balanced within each condition. Accordingly, the first and the second vignettes presented to each participant differed on all varied factors, plus complainant gender. For instance, if a participant first read a vignette featuring a female complainant who alleged a contemporary/non-penetrative physical contact offence, was appropriately questioned and with judicial intervention, then the second vignette would feature a male historical claim of attempted penile-oral penetration, inappropriately questioning without any judicial intervention. The offence of non-penetrative sexual abuse occurred either at school or at church.

Analyses of the experimental data yielded significant effects of victim age on the professionals’ perceptions of the simulated CCTV cross-examination in terms of quality, fairness, and impact on the evidence reliability and credibility. Interestingly, the question type impacted only the perceived quality and fairness of the cross-examination of the complainant, whereas the presence or absence of judicial intervention influenced only the perceived reliability and credibility of the complainant. Specifically, age-inappropriate questioning by the defence lawyer in the simulated CCTV cross-examination was rated significantly more unfair and aggressive, compared to age-appropriate questioning. Likewise, when the judge did not intervene in the questioning by the defence lawyer, the professionals perceived the complainant as significantly less reliable and credible.

In contrast with surveys and qualitative studies, experimental methods are more powerful research tools because they permit causal inferences, and identify cause and effect in the relationships between variables. By manipulating factors in controlled experimental conditions,
confounding features can be isolated and controlled through the random assignment of the participants to these conditions. Through causal inferences, in evaluating the use of alternate measures for complainants of child sexual abuse, the experimental method employed in Study 3 uncovered the importance of considering the age of a complainant, the appropriateness of the question form, and judicial interventions during witness examinations. The strength of the experimental findings of this study was derived from the scientific method applied in an online ‘laboratory’.

Discussion

To efficiently facilitate evidence-based practice, the evaluation of field practice requires rigorous and scientific methods. In the next part of the article, the strengths and contributions of the multi-study evaluation methods in the Royal Commission project are discussed in turn, followed by future research directions flowing from the review of this research.

Firstly, the research methods employed in the foregoing Royal Commission evaluation project may improve legal practice with complainants of child sexual abuse. The evaluation uncovered significant challenges and issues in current practice. The research also generated substantial recommendations for improvements in the ways in which alternate measures are implemented in Australian jurisdictions. In particular, improvements are needed to the quality of in court questioning of complainants of child sexual abuse.

Secondly, the integration of various methodologies in the project advanced the field of law reform evaluation research. Four distinct methodological categories applied were (a) stakeholder surveys and interviews, (b) retrospective analysis of archival documents, (c) comparisons with recommendations arising from prior social science and psychology research, and (d) experimental studies. This extensive methodological variety served to enhance the comprehensiveness of the evaluation of the uses of special measures and court practices in the field. Multiple methods provided a more efficient assessment of the complexity and intricacy of child sexual abuse cases and child witness characteristics. Any limitations of one

35 Royal Commission, above n. 20, Section 30.5.
methodology were mitigated by and integrated with the strengths of other methodologies through the commonalities they shared in the single research project.

Thirdly, methodological strengths emerged in the course of the research evaluation project. For instance in Study 1, a semi-structured interview method was used along with an inductive bottom-up approach. While these two approaches are often used in research studies, combining them in a single interview design yielded a wealth of findings. In particular, various practice characteristics were reported for the shared common topics and areas specified in the semi-structured questions, so the responses could be contrasted and discussed by professional group and by jurisdiction. Furthermore, with the inductive, bottom-up approach, the justice professionals’ suggestions and thoughts in response to the open-ended questions expanded the findings beyond the topics that had been outlined by the researchers. For instance, they provided suggestions to improve trial scheduling and case management to accelerate the effects of alternate measures for complainants of child sexual abuse.

Furthermore, the research design was strengthened by supplementing the interview data in Study 1 with survey responses in Study 2. In this way, confidence in and the generalisability of the interpretation of the research results was strengthened. In addition, the qualitative analyses of the responses to the open-ended questions in the quantitative survey data provided further evidence of professional development needs and suggestions for practice improvement. Another example where qualitative and quantitative methods were combined in a single study was the methodology of Study 16. Regarding the challenges posed by inconsistencies in the child complainants’ evidence of sexual abuse, these methods enabled the identification of different types of inconsistencies in central versus peripheral information and also categorised the chronological source of testimonial inconsistencies. Again, statistical analyses of the inconsistency categories by complainant age group and by jurisdiction provided invaluable evidence to evaluate practice.

Finally, another strength of the research project was inclusion of an experimental design, an innovative methodological initiative. The factors that influence the fairness and quality of CCTV cross-examination were examined by means of an online experiment, involving five groups of criminal justice professionals. Experimental designs allow causal inferences, thus the most powerful evidence for practice is available through this method.

**Limitations and future directions**

Although the Royal Commission project with 17 studies showcased a variety of methods for evidence-based practice in law and psychology, there were limitations offer future research opportunities in this field. Sampling methods in this area of research pose fundamental challenges. Not surprisingly, the sampling in the 17 studies did not include random sampling or other methodologies that guarantee more representativeness of the samples. The non-random and convenience samples achieved among the interviewed professionals, survey and experimental participants, were more limited than the random sample of cases included in the file review, and of the trials from which the trial transcripts were obtained, along with videotapes of pre-recorded police interviews and CCTV cross-examinations. However, this challenge is ubiquitous and an inherent feature of this type of research. Similar issues surfaced in a research project conducted by the UK Crown Prosecution Service (CPS) involving a review of 55 applications of special measures at the prosecution stage, yielding 10 recommendations for CPS.37 These data were gathered notwithstanding the absence of track indicators or monitoring flags for special measures within the CPS electronic systems.

Also, further methods to be explored beyond those reviewed in this paper may include the involvement of specialists and experts in research projects to evaluate evidence-based practice. For instance, a guide for British psychiatrists discussed the importance of special measures and fairness issues in procedures for children and vulnerable witnesses with psychiatric symptoms.38 The involvement of expert witnesses such as psychiatrists was recommended, along with ground rules hearings and liberal adjustments to court schedules to achieve more fair procedures and higher quality evidence. Focus group and survey methodologies are well-suited to canvass experts’ suggestions and recommendations about the practices of interest.

---

While a police service guide on working with vulnerable and intimidated witnesses has been available in the UK since 2011, the current evaluation provided guidance to police and other criminal justice sector professionals, with a focus on Australian courts and jurisdictions.\footnote{Ministry of Justice (United Kingdom), ‘Vulnerable and Intimidated Witnesses: A Police Service Guide’ (Report, March 2011).}

Another important source in evaluating the effectiveness of alternate measures for complainants of child sexual abuse is the complainants themselves and their caregivers. Their individual perceptions and views can be quantitatively analysed whereas their suggestions can be assessed by means of open-ended questions followed by qualitative analyses, similar to Study 2 above. Progress surveys and exit questionnaires built into the justice process will minimise the burden on complainants and keep the evaluations current. A substantial body of literature has established the significance of this research method for justice professionals, legal practitioners, policy-makers and court administrators.\footnote{Phoebe Bowden, Terese Henning and David Plater, ‘Balancing Fairness to Victims, Society and Defendants in the Cross-Examination of Vulnerable Witnesses: An Impossible Triangulation?’ (2014) 37(3) Melbourne University Law Review 539; Joyce Plotnikoff and Richard Woolfson, Evaluation of Young Witness Support: Examining the Impact of Witnesses and the Criminal Justice System (Home Office (United Kingdom), 2007); Elaine Wedlock and Jacki Tapley, What Works in Supporting Victims of Crime: A Rapid Evidence Assessment (Victims’ Commissioner for England and Wales, March 2016); Emmy Whitehead, 'Witness Satisfaction: Findings from the Witness Satisfaction Survey 2000' (Home Office Research Study 230, Home Office (United Kingdom), October 2001).}

**Conclusion**

Justice in cases of child sexual abuse is contingent upon eliciting evidence from the complainant. Yet, the psychological challenges and distress experienced by complainants during their involvement in the criminal justice process often compromise their evidence. To mitigate these effects, in Australia, various special and alternate measures for complainants of child sexual abuse have been implemented. The multiple research methods outlined in this paper demonstrate how these practices can be evaluated with scientific rigor and empirical sophistication. Together, the findings from the 17 studies indicate that alternative measures are supported by criminal justice sector professionals and are routinely used with child complainants. Nonetheless, five key areas were identified for improvement. These include the need to: (a) overcome technological obstacles; (b) align police interviews with evidence-based practice guidance; (c) improve the quality of in court questioning; (d) increase uses of alternate measures with adults; and (e) reduce delays in the prosecution process.
Empirical Methods in Law and Psychology: A Review Note on the Research Fostering the New Korean Jury System

Eunro Lee*, Yuhwa Han**, Kwangbai Park***, Sangjoon Kim**** and Minchi Kim*****

Abstract

Research related to the installation of the new Korean jury system in 2008 employed innovative empirical methods in the analysis of jury decision-making and its efficacy. The present research note describes various methods of a series of studies in the context of evidence-based practice, as the research findings were applied to the development of the Korean jury system. The methods of these studies in law and psychology harnessed the synergy and efficiencies of many multidisciplinary investigations. For example, information entropy theory, originating from physics and engineering sciences, was applied to understand jury deliberation, and a diversity index from biological research was adapted to conduct quantitative analysis of jurors’ interactions. Specifically, social network analyses, followed by information entropy comparisons, revealed how mock jury deliberations varied in terms of equality, richness, and evenness, depending on whether verdicts were bound by unanimity or majority decision rules. Probability theories and social and cognitive psychology theories were also applied and countered negative arguments often made regarding hung juries under the unanimity rule. Overall, the empirical methods proved to be powerful tools in supporting the design of the new system. Perhaps the methods can be applied to strengthening legal policy-making, establishing standards of judicial efficiency, and gauging the legal validity of jury outcomes in other jurisdictions. Further discussion is provided of the tensions and dynamics between the justice sector and the empirical science of law and psychology, surrounding the development of evidence-based practice in the Korean context.

* PhD, Lecturer, School of Health & Biomedical Sciences, College of Science, Engineering & Health, RMIT University, Australia. Author contact: eunro.lee@rmit.edu.au.
** Department of Psychology, Chungbuk National University, South Korea. Author contact: yuhwahan82@gmail.com.
*** Chungbuk National University, South Korea. Author contact: kwangbai@chungbuk.ac.kr.
**** Sangjoon Kim & Minseok Suh Law Office, South Korea. Author contact: ksjust@naver.com.
***** PhD, Department of Social Psychology, Sookmyung Women's University, South Korea. Author contact: mkim76@sookmyung.ac.kr.
Keywords: jury decision making, the Korean citizen-participation trial, methods of law and psychology, social network analysis, trend analysis, latent class analysis.

When Münsterberg¹, an experimental psychologist of the early twentieth century, was bringing research evidence into the law and the justice sector, the main use for experimental psychological research at the turn of the 20th Century was to provide evidence relevant to witness memory, false confessions, and related psychological factors. Since those days, empirical sciences have given legal professionals and theorists powerful research tools in an even broader range of legal practice and policy domains. The research inquiries often work on the micro level of trial and case evidence, for instance, with research topics such as face recognition and lie detection that help examinations in courtrooms. Furthermore, the research also sometimes supports macro level analysis inspiring legal system reform, such as research on structures of justice administration, and fundamental approaches to the exercise of the law.² Today, a growing body of multidisciplinary research assists law reformers achieve a better quality of justice across the globe.³

One notable area of collaborative multidisciplinary research has been developing in Korea, where legal scholars, justice sector officials, and behavioural scientists have joined to design and implement a jury system.⁴ This resonates with similar collaborative efforts made in Japan in the same period.⁵ Hans,⁶ as a comparative jury researcher, suggested that the Korean jury system (“Guk-min cham-yeo jae-pan”: citizen participation trials) provided a natural

³ David Canter and Rita Žukauskiene (eds), *Psychology and Law: Bridging the gap* (Routledge, 2008).
experimental laboratory for research concerning various jury systems in the world. This was because the new system was being ‘designed’ and was about to provide data from actual trials that quite unique research was stimulated. For the first time in their history, the people of the Korean peninsula now have a trial system involving lay participation. Launched in 2008, the Korean system has received both theoretical and practical scrutiny drawing on global and historical experiences and insights, particularly regarding its efficacy as a justice-delivery system in real-world circumstances.\(^7\)

Whereas in Europe the jury trial system traced back many hundreds, if not thousands of years\(^9\), the system signified a radical concept and social change for Koreans. The previous Korean criminal justice system was inquisitorial: trials were heard and decided by professional judges, and no lay citizens were involved. For cultural and historical reasons, popular acceptance of the inquisitorial legal system has been strong through the nation’s modern history. In comparison, the adversarial common law framework in the jury system was viewed with some suspicion. Consequently, the idea of adopting a system of jury trials—so-called lay citizen participation trials—met with significant scepticism and uncertainty. Critics pointed to the failings of jury trial systems in other countries and railed against the incongruity of an “alien” justice system implanted in their society.\(^10\) The official rationale finally given for the jury system relied on two arguments: (1) citizen participation in the legal system would help mature the nation’s democracy through providing the people with the power of practicing justice, and, (2) citizen participation would help restore legitimacy to the justice system that was seen to be untrustworthy at the time\(^11\). The public debate and speculations among justice professionals prompted scientific investigations with empirical research projects.

---


Responding to public uncertainty and scepticism, for instance, researchers in law and psychology were called upon to plan and execute sophisticated empirical studies to provide research evidence and bolster confidence about introducing a jury system in Korea. Even more experimental data were needed to design and implement the system.

The first major thrust of this research activity was in 2004, four years before the planned system was to commence. At that time, senior judges and social psychologists, working with members of the Korean Supreme Court, organized a series of mock jury trials as similar reform and research initiatives were being practiced in Japan.\textsuperscript{12} In addition, research funding was raised for a substantial project, encompassing 160 experimental mock trials, with approximately 1000 mock jurors in Korea. Furthermore 300 mock jurors in New York were engaged in the project to expand generalisability of the findings across legal cultures and jurisdictions. The results of those trials were documented in multiple academic research publications and public presentations\textsuperscript{13}. Further research was also completed after the jury system had been launched and utilized for actual criminal cases. For example, Park\textsuperscript{14} examined the actual verdicts from the new jury system and developed a measurement model concerning jurors’ ability and accuracy in their legal decision making. Park (2011) evaluated the fact-finding ability and accuracy of 1,318 juror eligible adults as fact-finders in criminal trials based on a sequence of their verdict decisions over several different trial vignettes. The mock jurors’ ability and accuracy were evaluated by the degree to which they applied the same implicit decision standard (“maximum alpha criterion”) consistently over the several different trials. From the study, it was found that the most accurate jurors with the highest level of fact-finding ability were also those who highly rated the importance of judicial instructions for their decision making. Given that a great portion of the judicial instructions consisted of explanations of due process principles and rules (i.e., presumption of innocence, prosecutor’s burden of proof, the standard of proof beyond a reasonable doubt, etc.), the


study found that the jurors with high ability and accuracy were those who apply due process principles and rules to decision making consistently over different trials.

The following sections describe empirical methods that played important roles in the development of the Korean jury system. Not all studies are reviewed because of the space limit of the present note and only a smaller group of example studies and their methods are discussed in relation to the new Korean jury system. For instance, even though the judge’s instructions to the jury explaining the standard of reasonable doubt has been investigated with a psycholinguistic research design, but it has been omitted in this review note.

Later sections of this paper review the design, launch, and implementation of the system, which span a period from 2008 to 2017. Though we focus on the empirical research during those developmental stages, we also address the tensions and dynamics between those generating the research evidence and those using the research findings for decision making, namely, the planners, policy makers, and other stakeholders. Finally, we suggest ways that the research methods related to the Korean jury system might also be applicable to judicial reforms in Australia and elsewhere.

**Research agenda for securing empirical evidence for designing a new Korean jury system**

The first and principal research topic preparatory to system design was to decide between the conventional jury system, which relies on a lay-citizen group of fact triers, or the Escabinado, i.e. a mixed jury system, which relies on both professional judges and citizen judges. The fact that Japan, a neighbouring country, was about to adopt a mixed-jury system in 2009\(^\text{16}\) had an effect on the debate in Korea.

A second research topic was to decide jury decision rule, one of the most controversial elements of jury systems worldwide.\(^\text{17}\) The options were the conventional unanimity rule and

\(^\text{15}\) Kim et al, above n 13.


a less stringent majority rule. Of particular concern was the fear of producing too many hung juries in delivering a verdict and inefficiencies in the system.¹⁸

Some research items were specific to the Korean context and sometimes widely to Asian civil law jurisdictions, for instance, the public attitudes of deference to authority that were debated in Japan as well¹⁹. Specifically, questions were raised about whether Korean jurors would deliberate competently and rationally as fact finders or whether they would yield to social pressure and conform to majority opinions. As mentioned previously, such concerns had historical and cultural origins. No tradition had prepared Koreans for the idea that lay citizens—their peers—could become legal decision makers. The conventional wisdom was that only judges—highly trained professionals—could be trusted as fact triers. Yet, ironically enough, it was the distrust and dysfunction of the old justice system that led to the call for lay citizens as jurors.

To clarify the best approach to take and decide upon an urgent research agenda, lawyers, policy makers, plus social and legal psychologists joined efforts to conduct mock trial studies and scientific experiments. In addition, a Presidential work force (the Presidential Committee for Judicial Reform, 2005-2006) was established and commissioned research and preparation tasks including the implementation of the new system. By those means, empirical approaches were brought to bear on the momentous decision to reform the criminal justice system for an entire nation.

Mock trials by the Korean Supreme Court

In 2004, the first two mock jury trials in Korea took place in a modified courtroom in the Central Seoul District. The trial scenario in both trials involved an alleged murder for which evidence of guilt was weak. Lawyers played the roles of the presiding judges, prosecutor, and defence lawyer. Professional actors played the roles of defendant and witnesses, including the victim. All jurors were volunteer citizens recruited from the local community. The two mock trials differed only in the type of the jury, with one being the conventional lay jury and the

---


¹⁹ Anderson et al, above n 5.
other being a mixed jury of judges and citizen jurors. The overall objective was to identify differences in jury behaviours and functions by experimental condition, i.e. type of the jury.

In the first trial, evidence was presented before three conventional juries of twelve persons each. After hearing evidence, all three jury panels deliberated and rendered a verdict of not guilty. The deliberations were video recorded, and the jurors responded to surveys before and after their jury service. In the second mock trial, evidence was presented before two versions of the European-style Escabindo mixed jury, one consisting of two judges and three jurors and the other consisting of one judge and five jurors. Again, jury deliberations were video recorded, and the jurors responded to surveys before and after their jury service. The analyses of data from the mock trials revealed significant differences in how the two jury types functioned. Those results, in turn, were critical to decisions regarding the general planning and design of the Korean system.

Understanding jury deliberation: Using trend analysis, network analysis, information entropy, and group psychology theories.

Trend analysis of jury deliberations. In Figure 1, empirical findings from the Korean Supreme Court mock trials are displayed. The results are from trend analyses over five phases of each jury deliberation. All 36 jurors from the three conventional juries participated in deliberations describing their thoughts and opinions, and the transcripts of their statements were analysed into units of individual utterances. Subsequently, each of those utterances were examined and coded for valence, that is, whether the utterance contained a positive, neutral, or negative term regarding the defendant’s guilt.

In the left chart of Panel (a), the five bars from left to right depict the valence distributions for each of five phases of jury deliberation. The majority of jurors began deliberations with the opinion “Not Guilty”, when they predominantly referred to the defendant in positive terms (more than 95% of the utterances). In the middle phases of deliberation, the jurors were more willing to speak negatively or neutrally of the defendant, with negative and neutral terms appearing in 40% of the statements (at Phase 3). Finally, the percentage of positive terms recovered in their utterances in the last phases, suggesting that by that time the jurors had established the notion of reasonable doubt into their deliberations. In the right graph of Panel

20 Park et al, above n 12.
(a) the U-shape curve plotted from the valence data is a graphic depiction of the trend in majority opinion during opposing arguments and the establishment of a reasonable doubt.

As shown in Panel (b) of Figure 1, the valence trend for the minority of jurors was starkly different. At the beginning of the deliberation, their utterances contained predominantly negative terms (70% at Phase 1). Yet, their consideration of positive arguments and terms regarding the defendant increased as deliberations continued. Finally, their utterances showed a higher proportion of neutral and positive terms for the defendant (around 50% at Phase 5). This upward trend, the J-Shape curve, indicated that they too had incorporated the principle of reasonable doubt into their deliberations. As it turned out, the minority group of jurors changed their initial opinion and came to agree with the majority in delivering a unanimous verdict of “Not Guilty”.

(a) The distribution of utterance valence among the majority jurors with pre-verdict opinion of “Not Guilty”.
(b) The distribution of utterance valence among the minority jurors with pre-verdict opinion of “Guilty” or “Undecided”.

Figure 1. Trend analysis of jury deliberation: The minority jurors’ J curve reflecting cognitive conversion based on the standard of beyond a reasonable doubt and the majority jurors’ U curve presenting their consideration of the minority jurors’ arguments and evidence. Based on the results of Park et al., 2005.

For the interpretation of the results, the researchers conducting the study drew upon a social psychological decision-making theory: conversion theory.22 These data clearly showed that the minority jurors did neither conformed immediately nor effortlessly to the opinions of the majority jurors. Rather genuine persuasion and opinion change in the opinions of minority jurors were achieved through discussing whether there was reasonable doubt about the defendant’s guilt. In addition, the ability of jurors to render unanimous verdicts through rational group processes was attributed to the openness of majority jurors to the opinions and arguments of minority jurors. Of further particular interest was the acquisition and display of experimental evidence to depict the discussion of reasonable doubt during jury deliberations.

The implications of the empirical approach in the study just described, which was tailored to meet the evidence needs for the jury system designing, was this: Korean lay persons were capable of serving as jurors and examining the evidence rather than merely agreeing too

---

easily with the majority view. They showed they could participate in rigorous group processes to arrive at credible and rational legal decisions. This crucial finding gave policy makers, think tanks, and the general Korean public the confidence that a jury system was an achievable goal23.

**Social network analysis and information entropy of jury deliberations.** Another remarkable method of social network analysis was used to investigate jury behaviour, as illustrated in Figure 2. In the diagram, the size of the dot/circle is proportional to the frequency of statements uttered by the respective juror during jury deliberation. Each line depicts the intervention path between any two jurors who spoke in response to or after each other’s statement. The thicker the line is, the more frequently the two jurors intervened, or communicated, with each other. In this manner, the social network diagram gives interesting graphic expression to the data from group deliberations. For the actual analysis of the data (including the frequency of each juror’s utterances, the number and direction of intervention paths, and the width of each intervention line) the researchers applied information theory24 and calculated a species diversity index.25

---


After they developed three descriptive indices with this type of entropy analysis, Park & Lee (2006)\textsuperscript{26} compared the equality, richness, and evenness of the three conventional-jury deliberations and those of the two mixed-jury deliberations. The results of the mixed-jury deliberations were substantially different from the results for the conventional-jury deliberations. The researchers, therefore, called for further studies to determine the effects of a wider range of factors, such as jury size, the judge’s style of leading deliberations in the mixed juries, and the percentage and characteristics of jurors in the subgroup representing the majority pre-verdict opinion, the subgroup representing the minority pre-verdict opinion, and the subgroup of undecided jurors. The results of the social network analysis were influential in the decision to adopt an adjusted conventional-jury system for the new Korean system, rather than the mixed-jury system being implemented in nearby Japan. The then-President task force was aided with this empirical evidence\textsuperscript{27} which highlights the power and importance of conducting tailored research in anticipating and advance of legal reforms.

**Laboratory mock trials and research evidence upon jury decision rules and hung juries**

In contrast to the Korean Supreme Court mock trial projects, further laboratory mock trials used either an edited trial video or a script vignette of a criminal trial. For the script vignette,

Figure 2. Network analysis of jury deliberation: The larger dots represent more statements by the jurors during the deliberation and thicker lines depict more frequent communication between the two jurors. Using this entropy analysis, equality of juror utterances, richness of interventions, and evenness of interventions were indexed and compared among different types of juries. Based on the results of Park & Lee, 2006.

---

\textsuperscript{26} Park and Lee, above n 21.

\textsuperscript{27} Kwangbai Park, Personal retrospections of the committee meetings and discussions.
the criminal case material was adjusted from that of a widely-used scenario. The video materials, however, were edited versions of the mock trial in the Korean Supreme Court project, or, in the case presented to the mock jurors in New York for expanded external validity of the research, they were edited video recordings of an actual American criminal trial. Details of the case being tried were manipulated to vary the strength of the evidence; there were four experimental conditions: (a) conflicting evidence, (b) strong evidence supporting innocence, (c) strong evidence supporting guilt, and (d) weak and vague evidence.

A second experimental factor involved the jury decision rule: each jury was randomly assigned either to the unanimity rule condition or to the majority rule condition. The jury size was chosen as eight persons as an even number close to the largest potential jury size (N = 9) to compare the verdicts between the unanimity and simple majority rule conditions.

Entropy analysis of the verdict data from 80 juries of eight persons showed that juries under the unanimity rule were more likely to render the same verdict for the same case (verdict stability) compared to juries under the majority rule. This greater verdict stability under the unanimity rule was related to deliberation characteristics. Especially for the cases with conflicting and vague evidence (experimental conditions (a) and (d) above), the jurors with majority opinions were more likely to consider arguments from jurors with minority opinions as the deliberation went on under the unanimity rule. Leniency contract theory was applied to interpret why the unanimity rule disposes majority jurors to open up to minority opinions.

**A measurement model of juror ability and confidence in the new Korean system**

Whereas the Japanese mixed-jury system revived in 2009 was motivated to reflect common sense and community values in criminal justice decision making, the motivation driving the

---

29 Lee et al, above n 13; Lee and Park, above n 13.
new Korean system was to extend to the lay citizen the power, rights, and obligations to engage in legal decision making, overcoming the lack of legitimacy of Korean judicial decision making. In other words, the Japanese mixed-jury (‘saiban-in’: jurors) system represented a model of cooperation between the professional judges and lay judges in a trial. In contrast, the Korean jury system placed lay jurors as independent decision makers on the defendant’s guilt. Although the un-binding effect of the jury verdict to the final trial verdict by the judge has been yet to be rectified, deliberations by jury and its rendering a verdict represent the independency of lay jurors’ legal decision making.

Accordingly, a critical question for the planning and design of the Korean jury system was whether lay jurors had the requisite skills and personal attributes to perform reliably as high-level fact finders. In answer to that question, Park developed a sophisticated model that shows the relationships between juror decision accuracy, juror ability, and trial difficulty. As a comprehensive psychometric approach, the Cronbach alpha reliability index and item response theory (IRT) analyses were utilised. Subsequently, the model showed that the ability of jurors may affect their accuracy, especially in difficult cases, more than do their attitudes, values, and biases. This research demonstrated that the abilities and decision accuracy of jurors can be analysed empirically, and that fact alone may have significance for similar research around the globe.

**Ongoing research and developments since system launch in 2008.**

The new Korean jury system (or the initial provisional version of the system) commenced in 2008; for the first time, lay citizens played significant roles as legal decision makers in the criminal justice system. The system emerged and was shaped from intense public debate and legal policy decisions by almost a decade as well as rigorous empirical research efforts.

---


33 Park, above n 14.

34 Hong Kyu Park, *Trial by Citizens!* (Saramsaenggak, 2000).
Finally, the new system was given a five-year trial period and amendment bills are still being debated at the time of this writing in 2017.

During the first phase of its provisional existence (2008 - 2012), the new system had three distinctive features. First, a jury’s verdict has not been binding due to constitutional reasons of the defendant’ right to be tried by a judge as well as the infancy of the confidence in the system. Instead, the lay verdict has been considered advisory for the presiding judge who delivers the actual verdict. This fact moved Han & Park to comment that “the Korean system should be called ‘trial with jury’ rather than ‘trial by jury’ that is found in common law countries such as the USA and Great Britain”. Second, the jury size has varied depending on the type of case and severity of the offence as in Japan. Third, jury decisions have been subject to the unanimity rule. However, if a jury fails to achieve a consensus after deliberating for some duration, that jury should consult the judge and render a verdict under a majority rule. Those three features and the research relevant to each are detailed in the following sections.

Advisory status of the jury verdict

The initial advisory status of the jury’s verdict and the stepwise introduction of the new system were concessions to prudence over the law reasons of the constitutional challenge from code of the trial by judge: they eased the transition from the familiar inquisitorial system to the new jury system. In addition, the transition period was a time to address public and official misgivings about the use of experimental findings to bring about major reforms. The advisory nature of the current jury verdict is fundamentally related to legal debates.

Specifically, the Korean Constitution Article 27(1) describes "all citizens shall have the right to be tried in conformity with the Act by judges qualified under the Constitution and the Act". Some proponents of the Korean jury system argue that “Judges qualified under the Constitution and the Act” can be interpreted as authoritative effects of jurors' verdict should be recognized. Contrastingly, the opponents insist limiting it as a trained professional judge’s

---

36 Han and Park, above n 4.
38 Park et al, above n 12; Park and Lee, above n 21; Park, above n 14.
39 The Constitution of the Republic of Korea, art 27(1).
because the Korean Constitution Article 103 specifies that "judges shall rule independently according to their conscience and in conformity both itself and other relevant laws".

The civil forum by the Korean Supreme Court (2013)\textsuperscript{40} also suggested actual binding of the jury verdict with the requirements of respecting and acknowledging the jury verdict for the trial verdict. Furthermore, the first clause of the Korean Constitution reading “All the power of the nation is from the citizens” will override the constitutional debate because jurors as citizens should be engaged in legal decision making, part of the judicial pillar of the three national powers (Korean Institute of Criminology, 2011)\textsuperscript{41}.

Although the advisory status of the jury verdict has been limited, the Korean jury system has been unique in that each real-world trial resulted in two verdicts, one from the jury and one from the judge(s). Researchers were quick to notice that a system of jury trials with two-verdict outcomes provided a rare opportunity to obtain real-world decision data with minimal measurement errors.\textsuperscript{42} It is because verdict data are available in real time rather than retrospectively from participant interviews and surveys, as in other studies.\textsuperscript{43}

Indeed, Kim, Park, Park, & Eom\textsuperscript{44} were able to replicate the classic research by Kalven Jr & Zeisel\textsuperscript{45} by using the real-world, real-time Korean trial data. In that study, the researchers compared 323 two-verdict sets from all Korean criminal jury trials from 2008 to 2010. Their analysis revealed that the judge-jury agreement was as high as 84% - 97.9%, depending on the strength of evidence. In the verdict pairs with disagreement, juries were more likely to render the verdict of not guilty, an outcome that was consistent with the jury leniency

\textsuperscript{40} The Supreme Court of the Republic of Korea, Media Release (18 February 2013) <https://blog.naver.com/1invincible/90166617841>
\textsuperscript{41} Korean Institute of Criminology, ‘Studies on the Criminal Justice Policies and Judicial Systems (V)-Focused on Evaluation Research on Civil Participation in Criminal Trials’ (2011)
\textsuperscript{43} Eisenberg et al, ibid; Kalven and Zeisel, ibid.
\textsuperscript{45} Kalven and Zeisel, above n 42.
hypothesis.\textsuperscript{46} Those results were strikingly similar to the findings of Kalven Jr & Zeisel\textsuperscript{47} and those of the relatively recent replication by Eisenberg et al.,\textsuperscript{48} despite the cultural differences between Western societies (with adversarial systems) and Korea (with the inquisitorial system).

Furthermore, empirical evidence on the jurors’ ability as decision-makers became obtainable. Kim et al.\textsuperscript{49} estimated and compared the accuracy of the juries’ and the judges’ verdicts by using latent class analysis. When the true verdict was guilty (that is, inferred true by statistical models), the probability that the verdict delivered by the judge was accurate was 1 and greater than that for the verdict delivered by the jury (probability = .88 - .97) depending on the models. More remarkably, the probability of a wrong conviction by the judge (probability = .07) was higher than that of the jury (probability = .02). Those results suggested that wrong-conviction rates would decrease when the jury verdicts became binding rather than advisory.

The power and strength of this method suggest a methodology for inferring true verdicts that can otherwise never be known. The researchers utilised latent class analysis whereby a latent (unobserved) variable, such as a class or a category, is estimated and identified for each case while the analysis focuses on an observed variable. The true verdict was inferred by estimating the probability of each verdict option (guilty or not guilty), while taking into account the actual verdicts by judges and juries and strength of the case evidence.

In light of those results, some Korean studies have gone on to argue that the jury verdict should be binding, that is, be officially recognized as the sole and final verdict of the trial.\textsuperscript{50}

Indeed, the Korean Supreme Court has established a precedent for the jury verdict:


\textsuperscript{47} Kalven and Zeisel, above n 42.

\textsuperscript{48} Eisenberg et al, above n 42.

\textsuperscript{49} Kim et al, above n 44.

“If the jurors rendered a unanimous verdict of acquittal after their attendance to the whole trial processes, based on the trial evidence including witness evidence, and the judge’s final verdict as her or his judgement was the same, then the verdict by the District Court should be preserved and not reversed by the High Court. The verdict by the District Court should be respected, unless new distinct evidence is provided to the extent to which High Court should initiate a new evidence hearing and subsequently finds the previous verdict inappropriate”.

This decision handed down by the Korean Supreme Court revealed the impact of empirical research evidence upon a relevant actual trial.

Another example of how the justice sector responded to research findings came from a committee in the Korean Justice Department that oversaw citizen participation in the justice system (hereafter cited as the Committee). In 2013, the Committee proposed an amendment to an act [Section 5 in Clause 46 in proposed Amendment of Act 11155 on Citizen Participation in Criminal Trials (hereafter cited as the Act)] regarding jury trials. The proposed amendment read as follows: “The judge should respect Items 2 and 3 (the jury’s verdict) when making a decision about the guilt of the defendant”. This proposed revision corresponded to the research findings by Kim et al. Unfortunately though, seven months later this item was further amended to render the jury’s verdict less binding than in the previous Act. More circumstances were added in which the judge was allowed to deliver a verdict different from the jury’s.

Jury size

Originally, the Act (Items 1 and 2, Section 1 in Clause 13) stipulated that the number of jurors depended on the case type. Specifically, nine jurors composed juries for case trials involving maximum sentences of capital punishment or life imprisonment; seven jurors for trials involving lesser criminal offences, and five jurors for trials in which there was little

---

53 Section 5 in Clause 46 in proposed amendment of Act 11155 on Citizen Participation in Criminal Trials, (2013).
54 Kim et al, above n 44.
ambiguity in the factual evidence. The rationale for this jury size categorisation scheme was to settle upon the most efficient jury sizes for the anticipated new jury system.\textsuperscript{55} In that way, the final legislation of jury size in the coming amendment would aim to achieve the best reflection of public values in legal decision making, provided the jury rendered verdict in consistency and with representativeness of the community.

As it turned out, five-person juries rarely served; only 6.8\% of the jury trials (N = 94) from 2008 to 2014 had five jurors. The average deliberation time was 1 hour 41 minutes, and the range was great: from 10 minutes to 5 hours 40 minutes. Deliberation time varied little across different types of offences. Although the 2015 report by the Court Office did not report on trial conditions, factors likely to have affected the duration of jury deliberation were complex and conflicting evidence, jury size, and other characteristics. The five-person juries spent an average of 1 hour 14 minutes in deliberation (with a maximum of 2 hours 40 minutes), which was similar to the average of 1 hour 30 minutes for seven- and nine-person juries. In 2013, the Committee proposed an amendment to use seven- and nine-person juries exclusively, thereby ending the use of five-person juries. This change of the jury size remains in the amendment proposal that has been announced for legislation in December in 2013.

\textit{The jury decision rule}

The Act (Section 2 in Clause 46) prescribes the unanimity rule for jury decision-making, but if the jury fails to reach a unanimous consensus for a verdict, they can then decide under the majority rule, but only after consulting the judge. In the period from 2008 to 2014, in 68.5\% (n = 1003) of the jury trials, the judge’s final verdict was the same as the jury’s unanimous verdict, whereas in 13.2\% (n = 194) of the trials the judge’s verdict was the same as the jury’s majority verdict.

Previous Korean experimental evidence suggested that verdicts under the unanimity rule showed superior stability than those under the majority rule.\textsuperscript{56} In addition, deliberations under the unanimity rule appeared to be more elaborate and rigorous, and jurors experienced greater satisfaction and had stronger confidence in their decision, compared with jurors deliberating under the majority rule. On the other hand, if the jury deliberating under the unanimity rule

\textsuperscript{55} National Court Administration, ‘Interpretation of “the Gukmin-eui Hyongsa Jaepan Chamyeo-e Gwanhan Beobryul [Act on Citizen Participation in Criminal Trials]” (2007).
\textsuperscript{56} Lee and Park, above n 13.
failed to reach a verdict and became a hung jury, the socioeconomic costs were high and the system felt inefficient.\(^\text{57}\)

Apparently, there have been case trials stymied by hung juries in which a majority rule would have brought a verdict.\(^\text{58}\) To ensure that the justice system could avail itself of the benefits of both types of decision rules, the Act incorporated unanimity as the default rule, with majority rule after consultation with the judge as the expedient rule if the jury failed to reach a unanimous consensus.

Initially, the majority rule referred to in the Act was defined as the requirement that a simple majority would rule. That is, the opinion with more votes (e.g., five votes in a nine-person jury) would become the final verdict. Elsewhere (in England, for example), definitions of *majority* have undergone refinements to produce supermajority rules. Such decision rules set the criterion for a majority. Thus, a final verdict, a majority, may often be set at two votes out of three or five votes out of six.\(^\text{59}\)

The simple majority rule initially embedded in the Act certainly expedited the delivery of a jury verdict, but at a cost. Jurors put less thought into their deliberations, gave lighter consideration to conflicting opinions, and made fewer attempts to persuade their peers.\(^\text{60}\)

Consequently, the Committee’s amendment proposed a supermajority rule in 2013, with the criterion for a majority set at three jurors out of four. This provision for supermajority rule retained the requirement that the jury consult the judge when a unanimous consensus became impossible.

This amendment proposal for a supermajority decision rule has been criticised by Lee (2014),\(^\text{61}\) who contends that the judge’s opinion imposes on the independence of the jury in

---


\(^\text{59}\) Devine et al, above n 47; Guerra et al, above n 59.

\(^\text{60}\) Devine et al, above n 47; Hans et al, above n 18; Lee and Park, above n 13.

its decision-making. Studies, moreover, have argued that the majority rule can diminish the democratic aspects of jury decision-making. Still other studies have complained that the rule contradicts the standard of beyond reasonable doubt for a conviction, because a minority opinion of not guilty goes unresolved.\(^{62}\)

The most controversial element in the proposal, however, was that it allowed judges to render verdicts according to their own judgment whenever juries could not reach a supermajority decision. In effect, this decisional setup was tantamount to a jury system without jury involvement. Kim (2016)\(^ {63}\) argued that redeliberation by the jury would provide better justice, and it is prescribed by the Austrian Criminal Procedure Act Clause 332 Item 4. In this alternative, if redeliberation in order to render a verdict is necessary, the jury remains the decision-making engine of the system.

**Amendment of the Act**

Since the Korean jury trials were introduced in 2008, scholars of law and empirical sciences have continuously sought ways to strengthen the jury’s impact on trial proceedings. The examples are proposals for supermajority rule for jury decision-making, the Committee’s amendment proposal, and the final amendment recommendation by the Korean Justice Department.\(^ {64}\) There has been some criticism that the fundamental changes contained in the Act do not push the justice system far enough beyond its initial version. Moreover, the amendment proposals, which were announced for legislation in March, October, and December of 2013, have yet to go into effect. The ongoing research and the empirical work already incorporated into the amendment bills are clear and present forces for salutary change, but they alone are not enough even when enacted. If the Korean jury trial system is to mature properly and reach its full potential, it will require the proactive support and full commitment of the public and governing bodies.

**Summary: Interactions between the empirical sciences and justice system to foster the new Korean system**

We now review the journey shared by empirical researchers, scholars in law and psychology, and public policy makers who have collaborated to foster the Korean jury system; a

---


63 Ibid.

64 See Lee, above n 61; Kim, above n 62.
profoundly complex undertaking. Globally, the journey began more than 100 years ago, when Münsterberg introduced experimental methods and research evidence to the American legal system. The journey continued when a number of Korean empirical scientists and legal reformists scientifically responded to the social pressure and desire for democratic reform in the justice sector since the 1990s, due to the public distrust and perceived unfairness in the judge trial system.  

In the early stages of the reform, the Korean Supreme Court conducted mock trials to weigh the merits of the conventional Anglo-American jury system and the Escabinado mixed-jury system of Europe and to identify possible factors peculiar to Korean jurors. As described previously, the trial locations and the participants were carefully chosen to ensure the ecological or external validity of those field experiments, a typical target of criticism regarding mock trials. Monitoring the trials were teams of experimental social psychologists and reformist justice professionals, including Supreme Court judges.

Content analysis techniques were applied to the deliberations of the three conventional juries and the two mixed juries in the mock trials. Those techniques quantified three key characteristics of jury deliberations: equality, richness, and evenness. In addition, the researchers developed descriptive indices by applying techniques from social network analysis, information entropy, and species diversity theories. The result of this massive research effort was the empirical research evidence supporting a jury system—one built along the lines of the Anglo-American system of conventional juries—was a feasible choice for the Korean system.

The experimental methods developed to compare the two types of jury proved to be not only innovative, both theoretically and practically, but also effective. Significantly, civic planners in other countries, for example, Taiwan, may find the experimental methods to be valuable tools for designing or reforming their own justice systems.

---

65 Han and Park, above n 4; Hans, above n 6.
66 Park and Lee, above n 21.
68 Huang and Lin, above n 47.
A fundamental concern about the installation of the new jury trial system was whether the average lay Korean had the temperament and wisdom to make grave and delicate legal decisions appropriately and fairly. This concern was predictable, since Korea had no historical or cultural precedent for such a system. To address this concern, researchers relied on the content analysis approach and trend analysis technique performed on juror deliberations in the mock trials. They found that the unanimous verdicts from three mock juries indicated that minority jurors could assimilate the reasonable doubt standard into their deliberations, so that they could eventually join with the majority opinion regarding not-guilty verdicts. This was empirical evidence of jurors ability to perform rationally in group decision-making. Furthermore, it refuted arguments that Korean jurors would bend to social conformity rather than decide the verdict independently and rationally. Finally, Park showed that juror accuracy, juror ability, and trial difficulty could be modelled empirically with psychometric precision.

Choice of another system design element—the jury decision rule—required more conventional laboratory experiments, following the methods and findings from the Korean Supreme Court mock trials. To meet the challenge, the researchers examined data from more than a hundred mock juries with more than a thousand mock jurors. In keeping with previous research regarding juries, the researchers settled on the superiority of the unanimity rule over the majority rule. Specifically, conditional probability theory showed that juries under the unanimity rule appeared to render more stable verdicts across mock juries. Specifically, multiple juries that served in the same mock trials were more likely to reach the same unanimous verdicts, whereas juries under a majority rule showed greater variability.

The use of the unanimity rule incurs the problem of the hung jury, which is both a curse and a blessing in a justice system. On one hand, of course, it is a costly waste of trial time and witness and lawyer energies due to the jury’s failure to render a verdict. On the other hand, as the Korean mock jury data demonstrates, hung juries deliberated more systematically, and they more closely scrutinized the trial evidence. Even though the minority jurors in hung juries felt more pressure to conform to their majority fellows, their comprehension of the

69 Park et al, above n 12.
70 Park, above n 14.
72 Lee and Park, above n 13.
evidence was not inferior to the minority jurors in the juries that delivered verdicts. The conclusion drawn from this research evidence was that a hung jury reflected the complexity of the case. Consequently, a retrial was not a waste of resources but rather a continuation of the due process of a fair trial.  

The launch of the Korean jury system encountered some delay. The public required time to get used to the idea of the jury trial, there were general doubts about its feasibility in Korea, and political resistance arose among anti-reformists within the justice sector. Nevertheless, the use of empirical research, finely tailored to political, social, and legal realities, gave the proponents the knowledge and research evidence they needed to win public support.

Today, the Korean jury system is still in its infancy, with an amendment bill having been presented since 2013. The Korean Parliament has yet to either pass or reject the bill. But this much is already certain: the collaboration of empirical scientists, social scholars, and political leaders has enabled Korea to achieve a magnificent goal, namely, a more democratic system of justice. For the first time, citizens have a role in an institution that preserves justice in their society. They have become participating stakeholders in the society that sustained them.

**Justice and the jury systems around the world: The search for best evidence-based practice**

Nolan & Goodman-Delahunty characterized the Australian criminal jury trial as an endangered species, with less than 1% proportion in the criminal cases, and further declared that civil jury trials were also on the verge of extinction. Nonetheless, presuming that juries provide social benefits, they called for further research on jury deliberation, not only at the level of the individual juror but also at the jury level. This call was to understand and improve jury dynamics and the legal decision processes. Such research would align precisely with the capabilities of the empirical methodologies developed for the new Korean system. For example, the use of mock juries or shadow juries in real trials would identify factors influencing verdicts; deliberation characteristics such as equality, richness, and evenness of jurors’ interventions; and social the influences and cognitive group processes among minority

---


74 Mark Nolan and Jane Goodman-Delahunty. Legal Psychology in Australia (Thomson Reuters, 2015).
and majority jurors. As further example, Lee et al. (2012)\textsuperscript{75} analysed data from 18 shadow juries and 16 research juries for a total of 20 criminal trials in Korea in the first half of 2012. The research methods adopted the information entropy analysis developed by Park & Lee (2006)\textsuperscript{76} and illustrate an application of the methodologies reviewed in the present note.

Recently, Horan & Israel (2016)\textsuperscript{77} stated that research with real juries should be harnessed to strengthen the Australian jury system. This suggestion was somewhat daring because, under the secrecy law, jurors are banned from sharing any experience or information regarding trial and deliberation processes they experience.\textsuperscript{78} Yet, there is much to learn about justice reform within the Australian legal system, including the benefits or drawbacks of timely and legitimate calls for evidence and the banning of hearsay. Here again, the research methods and collaborative approach of the Korean experience would have productive applications.\textsuperscript{79}

Globally, mixed juries and tribunals are often advocated,\textsuperscript{80} especially during the revival or introduction of jury systems in certain countries, including Russia, Mexico, Japan, and Taiwan.\textsuperscript{81} Because of the high regard for juries worldwide and the scrutiny they have received both in the professional discourses and in the press and fiction, the global norm seems to be the participation of lay citizens in the decision-making.\textsuperscript{82}

Even so, juries come in a variety of forms and sizes. For legal scholars and social psychologists, this miscellany of jury types throughout the world represents a bonanza of field research opportunities for the study of jury behaviours and dynamics.\textsuperscript{83} The possibilities for methodological innovation and collaborative research in law and the psychology of the

\textsuperscript{75} Jaehyup Lee et al, Exploring of improvement in jury deliberation with a shadow jury method. (The Korean Court Administration, Seoul, 2012).
\textsuperscript{76} Park & Lee, above n 21.
\textsuperscript{78} Jane Goodman-Delahunty and David Tait, ‘Lay Participation in Legal Decision Making in Australia and New Zealand: Jury Trials and Administrative Tribunals’ in Martin F. Kaplan and Ana M. Martín (eds), Understanding World Jury Systems Psychological Research (Psychology Press, 2006) 47.
\textsuperscript{79} Tait and Goodman-Delahunty, above n 37.
\textsuperscript{82} Mader and Hans, ibid.
\textsuperscript{83} Hans, above n 6.
courtroom are inexhaustible. Expected is that the techniques and the interdisciplinary approaches we have described here will shed a light on the way to discoveries for many years to come.

**Conclusion**

The jury, one of the most complex of human decision-making systems, has provoked enormous amounts of empirical research. At the macro level, the jury exemplifies our yearnings for democracy and even nullifies the law, if necessary. At the micro level, however, research brings empirical evidence, for instance, regarding illusions and bias in human vision, perception, and memory into courtrooms and jury rooms. In this era, human intelligence strives with the easy and hard problems of consciousness in the advance of quantum mechanics and free will is being examined at the neuronal level in the brain. Borders between disciplines are no longer clear and therefore we may predict law and psychology research will advance the jury system by application of scientific research methods drawn from many disciplines, and by utilising partnerships between actors able to synthesise normative study and empirical approaches. In this global context, the Korean jury system and the empirical evidence surrounding its birth and development may stimulate scientific interactions between practice and research efforts across the jurisdictions on the globe.
Using Pretest-Posttest Research Designs to Enhance Jury Decisions

Jane Goodman-Delahunty* and Natalie Martschuk**

Abstract:

When lay jurors are unfamiliar with key factual issues presented in criminal trials, the risk of wrongful acquittal or conviction can be avoided by providing educative information about these issues to assist juries in understanding the evidence. This study illustrates a systematic research program devised to increase jury knowledge about child sexual abuse (CSA). The effectiveness of an educative intervention in the form of specialised knowledge presented via expert evidence that targeted common jury misconceptions about CSA was tested in the context of a simulated videotrial. A pretest-posttest research design was applied to measure mock-jurors’ CSA knowledge gains attributable to the intervention and the extent to which accurate CSA knowledge predicted verdicts. In addition, individual juror decisions were compared with jury decisions following group deliberation. Results revealed that many jurors endorsed CSA misconceptions and that their knowledge predicted credibility assessments of the complainant and their verdicts. However, decisions by juries following group deliberations produced an unexpected increase in acquittals despite increased CSA knowledge after exposure to expert witness evidence. Applications of this research paradigm can assist courts and parties in determining the most effective method to use to enhance jury decisions in criminal cases involving specialised knowledge beyond the expertise of most lay jurors.

Keywords: child sexual abuse, jury deliberation, expert evidence, pretest-posttest research design

* Research Professor, Faculty of Business, Justice and Behavioural Sciences, Charles Sturt University. Author contact: jdelahunty@csu.edu.au.
** Research Associate, Charles Sturt University. Author contact: nmartschuk@csu.edu.au.
Introduction

Although a comparatively small proportion of criminal cases in Australia are tried before a jury, child sexual abuse (CSA) cases comprise a disproportionally high percentage amongst them. The Australian petit jury consists of a group of twelve lay citizens who are empanelled to determine whether, based on the evidence presented at trial, the accused is guilty as charged beyond reasonable doubt. At the close of the trial and before jury deliberations commence, the judge orally summarises the relevant evidence, the applicable law, and presents jury directions on how to apply the legal principles to the facts in evidence. Following their deliberations, which are conducted in private and are nondisclosable, the jury has to render a unanimous or majority (one dissenter) verdict on each count. The verdicts that juries reach are presumed to reflect interpretations and inferences from the evidence and coherent explanatory narrative accounts developed by juries during group deliberation from the facts presented at trial, although direct observations and measures of real deliberating juries are prohibited by law. Due to these restrictions, with a few rare exceptions, most research on jury processes and decision making has been conducted with simulated trials and mock juries.

Jurors’ individual attitudinal biases influence their decisions, however, these predispositions are expected to balance out in a randomly selected jury of twelve, and to be less influential on verdict than considerations of the evidence. This was confirmed by controlled experimental studies in the form of jury trial simulations showing that the evidence presented at trial predicted about 80% of verdict, while juror demographics, attitudes and biases exerted comparatively little or no influence. However, early studies were typically conducted with

mock jurors who did not deliberate to verdict as a group, i.e., the unit of analysis was the individual mock juror and not juries comprised of groups of mock jurors. In more recent research, greater emphasis has been placed on the deliberation process of juries as a group and the ways in which the collective wisdom of the group or the knowledge of individual jurors impacts the verdict.

**Factual controversies in cases of child sexual abuse**

When the key issues in evidence are outside of the scope of common jury knowledge, the potential for widespread misconceptions or misinterpretations of the evidence to influence the ultimate verdict reached by a group of lay jurors increases.\(^8\) In CSA cases, the child complainant is frequently the sole prosecution witness because corroborative evidence in form of an eyewitness, medical or forensic evidence, is typically lacking. Thus, the factual evidence in controversy in CSA trials is typically based on the word of one witness against that of another. In these trials, jurors have to assess the oral testimony of the child complainant, and to determine which of the parties is more credible.\(^9\) In accordance with the principle of *in dubio pro reo*, juries are loathe to rely on uncorroborated evidence of a complainant, especially a child complainant, and tend to acquit.\(^10\) Compared to other types of criminal cases, rates of acquittal in CSA cases are disproportionally high.\(^10\) Statistically, these cases pose a greater likelihood of wrongful acquittal, rather than wrongful conviction.

One consequence of the impunity enjoyed by predatory child sex offenders is persistent reoffending against the same or other potential victims.

In response to concerns about the impunity of child sexual offenders, in many countries around the globe, special commissions of inquiry have been established to evaluate policies and procedures applied in the criminal justice process that make it difficult to achieve justice in cases of child sexual abuse. For example, in 2015, in the United Kingdom, a special commission of inquiry was established to examine the issues of delay and attrition in cases of child sexual abuse.\(^8\) Ian Freckleton et al, *Expert Evidence and Criminal Jury Trials* (Oxford University Press, 2016).\(^8\)


statutory inquiry into public failings in CSA cases was initiated, and is anticipated to take at least five years. In Australia, a Royal Commission into Institutional Responses to Child Sexual Abuse was established in 2013, and issued its final report in 2017.\(^\text{11}\) This Commission implemented an evidence-based inquiry, and sponsored more than 100 research projects on related topics, including a study of jury decision making in cases of historical child sexual abuse. The focus of that study was the extent to which three particular forms of unfair prejudice (characterological prejudice, accumulation prejudice and factual conflation) arise in joint trials in which tendency evidence is presented.\(^\text{12}\) Nonetheless, that study provided insight into the relationship between individual versus group knowledge about CSA, and the extent to which jury group knowledge was related to the ultimate verdict.

**Interventions to Reduce Juror Bias**

Past surveys of community members, including jury eligible citizens in several countries, revealed that lay knowledge of CSA was poor, and the presence of a gap between research findings on CSA and what many lay jurors believed. Jury simulation studies of CSA cases disclosed a robust relationship between juror susceptibility to misconceptions about CSA and individual mock-jurors’ verdicts to acquit.\(^\text{13}\) In other words, the greater the incidence of misconceptions, the more likely jurors were to acquit the defendant, whereas jurors with more accurate CSA knowledge were more likely to render a verdict to convict.

In the criminal justice process, the collective deliberations of twelve jurors in a group about the trial evidence is presumed to operate as a legal safeguard against jury errors based on the cumulative common knowledge of those jurors to appropriately interpret the evidence admitted at trial. The random selection of a group of twelve jurors from different walks of life in the relevant community, with different types of education and life experience, is anticipated to balance social predispositions, attitudes and biases that individual jurors may bring to court. Thus, the deliberation process is expected to reduce the impact on verdicts of individual juror biases. However, when the key factual issues in the case are widely misunderstood, the benefit of group deliberation in reducing the influence of individual

---


\(^\text{13}\) Jane Goodman-Delahunt, Anne Cossins and Kate O’Brien*44*(2) *Australian and New Zealand Journal of Criminology* 196.
biases on verdicts may be constrained. Accordingly, in cases in which specialised knowledge is required to understand and appreciate the facts in evidence, the criminal justice system permits educative interventions which are intended to reduce juror reliance on common misconceptions and biases in assessing the evidence adduced at trial. The most common educative intervention is expert witness testimony; however specialised knowledge may also be presented to jurors in the form of judicial instructions or directions about a topic. For example, to alert jurors to factors which can influence the reliability of an eyewitness to a crime, expert witnesses may offer evidence about general memory processes and ways in which memory errors arise. The same or similar information may also be presented to the jury by the presiding judge in the form of special jury directions. Similar provisions have been implemented in CSA cases to narrow the gap between common misperceptions about CSA and research findings on the topic.¹⁴

Legislative provisions enabling educative interventions in CSA cases have been implemented in all Australian jurisdictions and in the Uniform Evidence Act.¹⁵ For example, expert evidence on child development and behaviour¹⁶ or judicial directions on these topics can be introduced. The content of these educative interventions is specialised knowledge that falls outside of the common sense and common experience of most jurors. Generally, the content of the specialised knowledge is derived from empirical findings on the counter-intuitive behaviour of CSA victims, children’s suggestibility in response to questioning by adults, and the reliability of their accounts.

The effectiveness of the different types of potential educative interventions was compared in a meta-analytical review of empirical literature conducted two decades ago.¹⁷ The results of these analyses revealed that the introduction of expert evidence yielded a small but


significant effect on verdict. This effect remained significant when the topic of the expert testimony was child witnesses. The effects on verdict were comparable between groups of mock jurors who deliberated and groups of nondeliberating mock jurors.

Arguably, educative information provided by an expert might be a more effective and preferable method of exposing juries to the relevant specialised knowledge than the same educative content presented by the presiding trial judge because the former is subjected to cross-examination, and juries have the option to accept or reject the expert evidence. By contrast, the same evidence presented by the judge requires the judge to endorse a particular psychological theory, and the jury is directed to follow it without the benefit of cross examination.  

Mock juror responses to these two types of educative interventions have been tested empirically. The present article uses the example of educative information provided by an expert witness to illustrate a general methodological approach to assess the effectiveness of such interventions in jury trials.

Research Questions

This article describes a comprehensive research design applied in the context of a larger project funded by the Australian Research Council to understand and improve jury decisions on CSA cases. The aims of this methodological exposition using some data drawn from that project is to isolate in relation to the CSA evidence the extent to which (a) pretrial individual juror biases influenced jury decisions; (b) educative information presented by an expert witness influenced juror and jury CSA knowledge; and (c) deliberation influenced the verdicts.

Comprehensive Research Design

The following sections describe the set of research procedures and the research design that were applied to address the complex problem of juror bias and jury decision-making in CSA trials. An innovation of the systematic research program was the comparison of decisions of individual jurors versus deliberating groups, using realistic trial materials and actual nonempanelled jurors as opposed to students or citizens, increasing the ecological validity of the findings. Although some meta-analyses showed few differences in juror decisions by

---

18 People v Wright 248 Cal Rptr 600 (1988), II C1 [7a].
19 Discovery Project 110103706 awarded to Jane Goodman-Delahunty and Anne Cossins, “Countering Misconceptions in Child Sexual Assault Cases with Expert Evidence and Judicial Directions”.
students versus nonstudents, in sentencing decisions, student mock-jurors were more lenient than their nonstudent counterparts. The extent to which student and nonstudent deliberations are comparable is unknown as studies comparing students and nonstudents did not include jury deliberation.

Participants
Participants were nonempanelled NSW jurors who participated in simulated CSA trials conducted on court premises in the greater Sydney area. Participation was in-person, followed by completion of an individual posttrial survey questionnaire. Participants were randomly assigned to render either an individual decision (nondeliberating jurors) or to deliberate for up to 90 minutes in groups of 8-12 persons (deliberating jurors) before rendering a verdict.

Trial Simulation Materials
The trial simulation materials based on actual CSA cases and described a case of intrafamilial CSA. The defendant was the grandfather of a female child complainant, aged 11 years at the time of the alleged offence. The allegation was a single penetrative offence without the use of force. The complainant disclosed the abuse immediately to her grandmother after she was asked what had happened.

The professionally-acted video-trial included opening and closing arguments of prosecution and defence, direct and cross-examination of the complainant, a corroborative witness (her grandmother) an expert witness (in some versions), and judicial summation. The duration of the mock-trial ranged between 45 and 55 minutes (with educative information from the expert witness).

The educative intervention was included in the simulated videotrial. Specifically, the psychological expert witness provided information about common CSA misconceptions including empirical findings on counter-intuitive behaviour of CSA victims, children’s suggestibility in response to questioning by adults, and the reliability of their accounts (about

---

22 The trial scripts used in the experimental study were written by Professor Annie Cossins, with input from the first author.
2,200 words). The expert informed the jury that: (1) delays in reporting the abuse and continued contact with the offender, particularly when the offender is close relative of the child, are common; (2) symptoms and behaviours of children who are sexually abused are diverse; (3) physical or forensic evidence of the abuse is generally absent; (4) children are generally reliable witnesses; and (3) incomplete or inconsistent accounts are common. The expert reported that the complainant’s behaviour was consistent with that of a sexually abused child, but did not state that the complainant had been abused.

Dependent measures
The survey consisted of a questionnaire that measured common CSA misconceptions,23 perceived credibility24 of the child complainant and (if applicable) psychological expert witness, perceived factual culpability of the defendant25 and verdict26. Further, participants provided their perceptions of the evidence and the trial in open-ended responses. Video-recordings of the jury deliberations were transcribed for content analyses.

Data Analyses
Quantitative data were subject to descriptive and inferential statistical analyses, and analyses were conducted to compare test scores between-and within-participants. Data provided by jurors following group deliberations violated the assumption of independence for inferential analyses. Therefore, multilevel analyses were used to take into account dependent data by allocating individual jurors’ data to jury groups. At level 1, individual scores are expressed as group means plus the individual deviation of the group means, presenting within-group level analyses. At level 2, variability of group means is modelled and expressed as a grand mean plus the deviation of the group means, presenting between-group level analyses.27 The

23 The CSA misconception items were categorized into three main topic groups: (a) evidentiary features of child sexual offences; (b) children’s responses to sexual abuse; and (c) children’s suggestibility and reliability. Participants indicated their agreement with each of the items on a 5-point rating scale (1=strongly disagree; 5=strongly agree). Jane Goodman-Delahunty, Natalie Martschuk and Anne Cossins, ‘What Australian Jurors Know And Do Not Know About Evidence Of Child Sexual Abuse’ (2017) 41 Criminal Law Journal 86.

24 The Witness Credibility Scale is a 20-item semantic differential scale with four factors: Likeability, Confidence, Trustworthiness, and Knowledgeability. Participants rated the complainant and if applicable the expert on each of the paired adjectives using a 10-point rating scale, e.g. from 1=dishonest to 10=honest; from 1=illogical to 10=logical. Stanley L Brodsky, Michael P Griffin and Robert J Cramer, ‘The Witness Credibility Scale: An Outcome Measure for Expert Witness Research’ (2010) 28(6) Behavioral Sciences and the Law 892.

25 Participants indicated the likelihood that the complainant was sexually abused by her grandfather (1=very unlikely; 5=very likely).

26 Guilty vs not guilty.

27 Christian Geiser, Data Analysis with Mplus (Guilford Press, 2013).
advantage of these analyses is that they provide an indication of the extent to which the outcome is attributable to jury groups as opposed to the experimental intervention.

**Pretest-Posttest Research Design**

The study design was based on the randomised Solomon four-group pretest-posttest research design (see Table 1). This design involves two experimental (intervention) groups (Groups 1 and 3 in Table 1) and two control (no intervention) groups (Groups 2 and 4 in Table 1). All four groups provide posttest measures, whereas only Groups 1 and 3 in Table 1 provide pretest measures. This methodological approach allows a statistical comparison of posttest measures, the computation of differences between pretest and posttest measures, while controlling for pretest effects, and their interaction. More specifically, results of the Intervention groups are compared with those of the No intervention (or control) groups. To assess the influence of juror pretrial biases, both pretest and posttest measures are gathered from an Intervention and a No Intervention control group (Groups 1 and 2 in Table 1). To assess as to whether the experimental procedure of taking the pretest measure influences the posttest measure, for example, by sensitising the mock jurors to these issues, only posttest measures are collected from Groups 3 and 4, as shown in Table 1. In sum, this design allows tests of the magnitude of the effects attributable to the pretesting procedure separate and apart from the intervention.30

**Table 1. Solomon four-group pretest-posttest research design**

<table>
<thead>
<tr>
<th>Experimental group</th>
<th>Pretest measure</th>
<th>Intervention</th>
<th>Posttest measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>✓</td>
</tr>
</tbody>
</table>


An important feature of this design is the random assignment of participants to each experimental group (as opposed to procedures followed in a quasi-experimental design which may be similar without random assignment of participants to experimental groups). Random assignment ensures that each participant is allocated to one of the experimental groups at the same probability. This procedure decreases the likelihood of systematic differences between the experimental groups and thus avoids various types of selection biases. Consequently, any differences observed in the results between the groups can be attributed to the experimental manipulation, in this example, the expert witness evidence.

The advantages of this systematic research design include reduced error variance, that is, a reduction in the proportion of the variance in the results that is attributable to extraneous factors or measurement imprecision and not to the independent variable.\(^{31}\) Other advantages are increased internal validity, and the possibility to test causation. Internal validity is the extent to which an experiment is free from flaws in its structure and the results can be attributed to the true nature of the phenomenon in issue.\(^{32}\) This affects the reliability of the results obtained under controlled conditions with respect to the cause-and-effect relationships among variables.

**Assessing Juror Knowledge and Biases**

First step before testing the interventions is to develop a valid measure of juror biases. Therefore, a questionnaire was developed to assess jurors’ CSA misconceptions. A series of questions was drafted based on a comprehensive literature research to identify the most commonly endorsed CSA misconceptions and the extent of discrepancies between the knowledge of CSA experts and laypersons. The selected set of positively and negatively worded questions was administered to a group of non-empanelled jurors, and an item analysis was conducted to identify more precisely what jurors know and do not know about CSA.\(^{33}\) The results of these item analyses revealed that positively worded items were more robust than negatively worded counterpart items, due to the instability of the latter in conjunction with the positive and negative poles of the rating scale. Negatively worded items appeared to confuse participants in using the scale, resulting in within-participant response


inconsistencies and more extreme variation in responses.\textsuperscript{34} From the original set of 26 items 18 were used to assess their respective difficulty level. Table 2 displays examples of items of different difficulty levels that were included in the questionnaire. Items were identified as easy when a substantial proportion of eligible jurors (55\%) correctly responded to an item, and difficult when 45\% or more expressed uncertainty or incorrectly responded to an item.

Table 2. \textit{Pretest item analysis}

<table>
<thead>
<tr>
<th>Misconception</th>
<th>Uncertain</th>
<th>Error</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children who are sexually abused display strong emotional reactions afterwards.</td>
<td>33.5</td>
<td>44.8</td>
<td>78.3</td>
</tr>
<tr>
<td>Children aged 7-8 years are easily manipulated to give false reports of sexual abuse.</td>
<td>40.0</td>
<td>25.4</td>
<td>65.4</td>
</tr>
<tr>
<td>A medical examination almost always shows whether or not a child was sexually abused.</td>
<td>26.1</td>
<td>27.2</td>
<td>53.3</td>
</tr>
<tr>
<td>A sexually abused child typically cries out for help and tries to escape.</td>
<td>26.1</td>
<td>19.5</td>
<td>45.6</td>
</tr>
<tr>
<td>Children who change their reports of sexual abuse were probably lying in the first place.</td>
<td>25.9</td>
<td>9.7</td>
<td>36.6</td>
</tr>
<tr>
<td>A child who has been sexually abused will tell someone soon afterwards.</td>
<td>13.7</td>
<td>7.7</td>
<td>21.4</td>
</tr>
</tbody>
</table>

\textit{Note.} Agreement denotes greater susceptibility to CSA misconceptions.

The nine most difficult items, namely those with a combined rate of uncertainty and error rate that exceeded 45\% or more, were retained for validation analyses. Specifically, multivariate statistical analyses were conducted to test the factor structure (factor analysis) and predictive validity of the questionnaire.\textsuperscript{35} The factor analysis yielded a two-factor structure of the

\textsuperscript{34} Ibid.

\textsuperscript{35} For all further analyses the item values were reversed so that a higher number indicated greater knowledge about CSA, thus lower level of CSA misconceptions. Jane Goodman-Delahunt, Natalie Martschuk and Anne Cossins ‘Validation of the Child Sexual Abuse Knowledge Questionnaire’ (2017) 23(4) \textit{Psychology, Crime and Law} 391.
questionnaire: Factor 1 represented the Impact of Sexual Assault on Children, including the psychological and physical consequences or lack thereof (five items), Factor 2 represented the Contextual Influences on Report (four items). Further analyses revealed the predictive validity of the questionnaire: the fewer the juror misconceptions about CSA, the more likely jurors were to perceive the child complainant as credible, and the more likely they were to convict the defendant. The final nine-item questionnaire was used in a series of experimentally controlled interventions aimed at increasing jurors’ knowledge about CSA.

**Measures to Improve Juror Knowledge**

The following example applied a more complex variation of the Solomon four-group pretest-posttest design, in which the effectiveness of (1) group deliberation and (2) educative intervention provided by a psychological expert on juror decisions was tested. Table 2 shows the different experimental groups, mean pretest and posttest CSA knowledge scores, and juror verdict.

Table 2. *Juror CSA knowledge and verdict following (a) jury deliberations and (b) expert evidence*

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>$N$</th>
<th>$M_{\text{Pretest}}$</th>
<th>$M_{\text{Posttest}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. Expert evidence, no deliberation</td>
<td>108</td>
<td>28.1</td>
<td>31.0</td>
</tr>
<tr>
<td>1b. Expert evidence, deliberation</td>
<td>109</td>
<td>28.7</td>
<td>31.2</td>
</tr>
<tr>
<td>2a. No educative intervention, no deliberation</td>
<td>113</td>
<td>28.0</td>
<td>27.9</td>
</tr>
<tr>
<td>2b. No educative intervention, deliberation</td>
<td>109</td>
<td>28.0</td>
<td>29.1</td>
</tr>
<tr>
<td>3. Expert evidence, no deliberation, posttest only</td>
<td>231</td>
<td>--</td>
<td>31.1</td>
</tr>
</tbody>
</table>

*Note. Higher values indicate greater CSA knowledge. Mean score range: 9-45.*

Groups 2a and 2b in Table 2 viewed the mock-trial without educative information, whereas Groups 1a, 1b, and 3 were given educative information in form of expert evidence.
Participants in Groups 1b and 2b were assigned to jury deliberations, consisting of groups of 8-12 persons (dependent on participant availability on the day). Group 2a was a control group without any intervention or deliberation and provided the baseline comparison for both manipulations. Finally, Group 3 provided their CSA misconceptions scores only posttrial; the remaining four experimental groups provided this information in the pre- and posttrial questionnaires. The value of Group 3 in Table 2 was to test whether the pretrial questionnaire influenced juror responses posttrial. Although a fully crossed four-group pretest-posttest design includes a control group without any intervention and a posttest measure only (Group 4 in Table 1), we omitted this control condition due to logical difficulties (winter court recess without jury trials) and high costs of extending data collection to add a further 100 participants from a limited pool of non-empanelled jurors.

Despite the absence of Group 4 in our research design, results revealed no practice effect of the questionnaire when comparing posttest scores in the Groups 1a and 3 (Table 2): The CSA knowledge scores were significantly higher following educative information about CSA presented by an experimental psychologist regardless of exposure to the pretrial CSA questionnaire (equal posttrial scores) compared with the control group (Group 2a in Table 2). Furthermore, jurors’ CSA knowledge significantly increased following educative information about CSA (Group 1a in Table 2), whereas the knowledge scores remained unchanged in the control group (Group 2a in Table 2). Finally, the findings demonstrated that educative information provided by a psychological expert was more effective in reducing CSA misconceptions than group deliberations alone, reinforcing the view that key elements of CSA cases are not widely understood or common knowledge, thus need more attention in the criminal justice process.

---

36 One-way analysis of variance (ANOVA) was conducted to test the influence of education and questionnaire exposure on posttest CSA knowledge scores. Results revealed significant difference between the experimental groups, $F(2, 424) = 15.19, p < .001$. Post-hoc analyses revealed that significantly higher posttest CSA knowledge score following educative intervention regardless of pretrial exposure: Group 1a vs Group 2a: $M_{diff} = 3.2, p < .001$; Group 3 vs Group 2a: $M_{diff} = 3.3, p < .001$; Group 1a vs Group 3: $M_{diff} = -0.1, p < .995$. Because there was no practice effect of questionnaire exposure, Group 3 in Table 2 was ignored for further analyses.

37 Mixed between-within participants ANOVA, time of measure x educative intervention: Wilks’ Lambda = .88, $F(1, 204) = 27.38, p < .001, \eta^2_{\text{partial}} = .12$.

38 Mixed between-within participants ANOVA, time of measure x educative intervention: Wilks’ Lambda = .93, $F(1, 410) = 31.79, p < .001, \eta^2_{\text{partial}} = .08$; Mixed between-within participants ANOVA, time of measure x decision type: Wilks’ Lambda = .99, $F(1, 410) = 0.62, p = .433, \eta^2_{\text{partial}} = .00$. 
Assessing the Influence of Knowledge Gain on Juror Decisions

The finding that the expert witness successfully improved jurors’ CSA knowledge is important in determining the best practice and policy to educate lay jurors who serve on CSA trials to avoid decisions based on misinformation. However, it is also important to understand how this increased knowledge affects the perceived credibility of the complainants and ultimately, the jury verdict. Results revealed that the increase in CSA knowledge following the educative intervention did not produce a parallel increase in the perceived credibility of the complainant and verdicts to convict, as shown in Table 3.

Table 3. Perceived Complainant Credibility and Individual Mock-Juror Verdicts

<table>
<thead>
<tr>
<th>Experimental group</th>
<th>N</th>
<th>Mean credibility</th>
<th>Guilty verdict (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. Expert evidence, no deliberation</td>
<td>108</td>
<td>117.6</td>
<td>45.8</td>
</tr>
<tr>
<td>1b. Expert evidence, deliberation</td>
<td>109</td>
<td>120.2</td>
<td>25.7</td>
</tr>
<tr>
<td>2a. No educative intervention, no deliberation</td>
<td>113</td>
<td>114.5</td>
<td>39.3</td>
</tr>
<tr>
<td>2b. No educative intervention, deliberation</td>
<td>109</td>
<td>123.4</td>
<td>43.5</td>
</tr>
</tbody>
</table>

Inspection of the conviction rates disclosed an interaction between the educative intervention and the type of decision, i.e., whether the verdict was rendered individually by each mock-juror or as a group following jury deliberation. Whereas the conviction rate among nondeliberating jurors (Control group: 39.3%; Intervention group: 45.8%\(^{39}\)) increased somewhat, but not significantly, following exposure to the expert witness, it dropped sharply among deliberating jurors following exposure to the expert witness (Control group: 43.5%; Intervention group: 25.7%\(^{40}\)). The fact that the conviction rates in groups 1a (45.8%) and 2b (43.5%) were statistically undifferentiated indicated that the process of deliberation per se

\(^{39}\) \(\chi^2 (219) = 0.95, p = .330, \Phi = .066.\)

\(^{40}\) \(\chi^2 (217) = 7.63, p = .006, \Phi = -.187.\)
was not responsible for a decrease in conviction rates. In other words, these analyses provided no support for the leniency asymmetry bias\(^{41}\) or liberation hypotheses\(^{42}\) about deliberation effects.

Nonetheless, despite the improvement in mock jurors’ CSA knowledge following the educative intervention, that increase did not translate into more verdicts to convict the defendant. This research design showed that the unexpected sharp decrease in convictions in Group 1b compared to Group 1a was most likely attributable to influences during jury group deliberations. Thus further analyses of the jury groups as units were required to better understand this unexpected finding.

**Assessing the Influence of Groups on Decision-Making**

As was explained above, individual juror data violate the assumption of independency which is critical to conduct inferential and multivariate statistical analyses. Thus, analyses which take jury groups (as opposed to jurors) into account as units of analyses are necessary. Table 3 shows the mean pretest and posttest CSA knowledge scores, perceived complainant credibility and verdict for each deliberating jury.

Inspection of mean scores of the jury groups indicated large variability within the experimental conditions. For instance, the mean posttrial knowledge score ranged between \(M = 26.8\) and \(M = 32.9\) in the control group and between \(M = 27.8\) and \(M = 35.2\) in the intervention group. The variability was larger for perceived complainant credibility, ranging between \(M = 104.3\) and \(M = 134.6\) in the control group and between \(M = 102.7\) and \(M = 141.5\) in the intervention group. On average, the posttrial mean score of the intervention group (\(M = 31.2\)) was significantly higher than the posttrial mean score of the control group (\(M = 29.2\))\(^{43}\), whereas the difference was not significant for the pretrial CSA knowledge score (\(M = 28.7\) and \(M = 28.1\), respectively)\(^{44}\) and for perceived complainant credibility (\(M = 121.3\) and \(M = 123.4\), respectively).\(^{45}\)


\(^{43}\)\(t(19) = 2.20, p = .041\).

\(^{44}\)\(t(19) = 0.93, p = .366\).

\(^{45}\)\(t(19) = -0.40, p = .692\).
Further, the highest posttrial CSA knowledge scores were not necessarily associated with the highest complainant credibility ratings and proportion of guilty verdicts (e.g., Jury 4 in the intervention group, and Jury 7 in the experimental group, Table 3). By contrast, the highest perceived complainant credibility scores were associated with guilty verdicts (i.e., Juries 3, 5, and 8 in the intervention group, and Juries 3, 4, and 5 in the control group). These results indicated that the association between pretrial CSA knowledge score and verdict was mediated by the perceived credibility of the child complainant, supporting previous research findings with individual mock-jurors.46

A comparison in Table 3 of the individual mock-juror verdicts within each jury disclosed more hung juries among those who deliberated without the benefit of the educative expert evidence (Juries 1, 4, and 11) than among juries exposed to the expert evidence (Jury 4). The educative expert evidence appeared to assist juries in reaching consensus.

**Future analyses**

A further consideration to explore in future, for instance, is the relationship between the CSA knowledge score of the dominant person in each jury and the jury verdict. Although it can be argued that the foreperson is likely to lead the deliberation process, the foreperson is not necessarily the most dominant person in a deliberation, i.e., the person who speaks the most frequently may lead the discussion and influence other jurors the most. These analyses can be supplemented by in-depth content analysis of the transcriptions of jury deliberations to provide more detail about the impact of the expert evidence on group decision processes, by analysing the following: (1) deliberation topics (time and number of utterances); (2) endorsement or resistance to specific common CSA misconceptions (e.g., delay in reporting, suggestibility and reliability of children, post-abuse behaviour and reactions to sexual abuse, relationship with the abuser); and (3) assessments of the complainant’s credibility.

46 Jane Goodman-Delahunt, Anna Cossins and Kate O’Brien’A Comparison of Expert Evidence and Judicial Directions to Counter Misconceptions in Child Sexual Abuse Trials’ (2011) 44(2) *Australian and New Zealand Journal of Criminology* 196.
Table 3. *Mean Pre- and Posttrial CSA Knowledge, Perceived Complainant Credibility, and Verdict by Deliberating Jury Groups.*

<table>
<thead>
<tr>
<th>Intervention Group (1b) Expert evidence</th>
<th>Control Group (2b) No expert evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretrial CSA Knowledge</td>
<td>Pretrial CSA Knowledge</td>
</tr>
<tr>
<td>Posttrial CSA Knowledge</td>
<td>Posttrial CSA Knowledge</td>
</tr>
<tr>
<td>Complainant Credibility</td>
<td>Complainant Credibility</td>
</tr>
<tr>
<td>Verdict</td>
<td>Verdict</td>
</tr>
<tr>
<td>Jury 1</td>
<td>Jury 1</td>
</tr>
<tr>
<td>25.3</td>
<td>27.8</td>
</tr>
<tr>
<td>27.8</td>
<td>27.8</td>
</tr>
<tr>
<td>102.5</td>
<td>128.2</td>
</tr>
<tr>
<td>12NG</td>
<td>6NG</td>
</tr>
<tr>
<td>Jury 2</td>
<td>Jury 2</td>
</tr>
<tr>
<td>29.7</td>
<td>31.5</td>
</tr>
<tr>
<td>31.5</td>
<td>29.7</td>
</tr>
<tr>
<td>113.3</td>
<td>104.3</td>
</tr>
<tr>
<td>12NG</td>
<td>10NG</td>
</tr>
<tr>
<td>Jury 3</td>
<td>Jury 3</td>
</tr>
<tr>
<td>29.9</td>
<td>33.5</td>
</tr>
<tr>
<td>33.5</td>
<td>28.9</td>
</tr>
<tr>
<td>135.9</td>
<td>134.6</td>
</tr>
<tr>
<td>11G</td>
<td>8G</td>
</tr>
<tr>
<td>Jury 4</td>
<td>Jury 4</td>
</tr>
<tr>
<td>31.8</td>
<td>33.1</td>
</tr>
<tr>
<td>33.1</td>
<td>30.1</td>
</tr>
<tr>
<td>120.8</td>
<td>130.1</td>
</tr>
<tr>
<td>9NG 3G</td>
<td>7G 2NG</td>
</tr>
<tr>
<td>Jury 5</td>
<td>Jury 5</td>
</tr>
<tr>
<td>30.1</td>
<td>32.1</td>
</tr>
<tr>
<td>32.1</td>
<td>28.1</td>
</tr>
<tr>
<td>141.5</td>
<td>134.6</td>
</tr>
<tr>
<td>7G 1NG</td>
<td>12G</td>
</tr>
<tr>
<td>Jury 6</td>
<td>Jury 6</td>
</tr>
<tr>
<td>28.1</td>
<td>31.3</td>
</tr>
<tr>
<td>31.3</td>
<td>27.4</td>
</tr>
<tr>
<td>122.6</td>
<td>124.7</td>
</tr>
<tr>
<td>8NG</td>
<td>11G 1NG</td>
</tr>
<tr>
<td>Jury 7</td>
<td>Jury 7</td>
</tr>
<tr>
<td>26.4</td>
<td>28.1</td>
</tr>
<tr>
<td>28.1</td>
<td>29.6</td>
</tr>
<tr>
<td>105.9</td>
<td>122.5</td>
</tr>
<tr>
<td>12NG</td>
<td>9NG 1G</td>
</tr>
<tr>
<td>Jury 8</td>
<td>Jury 8</td>
</tr>
<tr>
<td>29.8</td>
<td>35.2</td>
</tr>
<tr>
<td>35.2</td>
<td>29.1</td>
</tr>
<tr>
<td>140.4</td>
<td>121.4</td>
</tr>
<tr>
<td>11G 1NG</td>
<td>9NG 1G</td>
</tr>
<tr>
<td>Jury 9</td>
<td>Jury 9</td>
</tr>
<tr>
<td>29.4</td>
<td>30.7</td>
</tr>
<tr>
<td>30.7</td>
<td>27.4</td>
</tr>
<tr>
<td>102.7</td>
<td>115.6</td>
</tr>
<tr>
<td>12NG</td>
<td>7NG 1G</td>
</tr>
<tr>
<td>Jury 10</td>
<td>Jury 10</td>
</tr>
<tr>
<td>27.3</td>
<td>29.1</td>
</tr>
<tr>
<td>29.1</td>
<td>27.0</td>
</tr>
<tr>
<td>127.5</td>
<td>127.9</td>
</tr>
<tr>
<td>10NG</td>
<td>9NG</td>
</tr>
<tr>
<td>Jury 11</td>
<td>Jury 11</td>
</tr>
<tr>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Note. G = guilty, NG = not guilty.
Conclusions

The research approach outlined in the foregoing example contributes to psycholegal scientific research on both CSA and jury decision-making. These methods extended past research by exploring reasons for the unexpected disjuncture between individual jurors’ CSA knowledge and jury group verdicts in CSA cases. The research advanced understanding of jury verdicts in CSA trials by explaining why a disproportionately high percentage of cases result in acquittals and identified legal policy reforms and interventions that can increase justice for future CSA victims. Results showed that the information provided by the expert witness improved juror knowledge about CSA and assisted the juries in reaching consensus. However, improved knowledge alone did not decrease the acquittal rate. Thus further research is needed before implications for policy can be provided. For instance, expert evidence that is more tailored closely to address factual issues specific to the case at hand, rather than generic educational information about CSA, and some guidance on deliberation processes may need to be tested.

By using a pretest-posttest trial simulation paradigm, changes in juror knowledge attributable to the interventions about key trial issues were uncovered. This new knowledge had an influence on jury appraisals of the credibility of the child complainant and on jury consensus. By including group deliberations, juror decision processes were uncovered and the importance of including deliberations in the research design was demonstrated. These insights are useful before making determinations about the effectiveness of interventions such as expert evidence about CSA on jury decision making and before providing recommendations to legal practitioners and policy makers about the effectiveness of these interventions.

---

Open Data: Turning Data into Information, and Information into Insights that Allow for Evidence-Based Policy

Alana Maurushat,* Jan Trzaskowski** and Hadeel Al-Alosi***

Abstract

This paper summarises the results of research into the legislative, policy, regulatory and operational enablers utilised in selected advanced jurisdictions (UK, US, France, Canada, New Zealand, the Netherlands and Denmark) as identified in the Open Data and Open Government Indexes which promote Open Data, a culture of data sharing and that can help inform future strategic developments in other jurisdictions. Open data presents the opportunity for industry, researchers and government to use previously privately held datasets to run analytics, turn data to information and information into insights that allow for evidence base policy. This paper analyses best practices from leading jurisdictions including analysis on legislation, responsibility and coordination, policies, regulatory settings, and operating environments as well as discusses exemplary open data projects, outcomes and applications.

Key words: open data, open government indexes, evidence-based policy

Introduction

The open government data movement is underpinned by the desire to make governments transparent, accountable and more efficient. The data driven economy relies on open data being machine readable and linked to allow advanced analytics and innovative applications.

---

* Alana Maurushat, Professor of Cybersecurity and Behaviour, Western Sydney University, author contact: alanacybersecurity@gmail.com. The research for this paper stems from commissioned research for the New South Wales Information and Privacy Commissioner contained in the report, Technical Country Reports on Enabling an Open Data Culture 2017 (unpublished). This paper has been re-written in an abbreviated format and contains new insights and findings. The findings of this paper reflect the opinion of the lead author and project leader Alana Maurushat. The Denmark section was written by Jan Trzaskowski and the Netherlands section by Hadeel Al-Alosi. The research from the other countries was as follows: UK by David Vaile and Alana Maurushat; US by Frank Smith and Alana Maurushat; France by Othmane Mechette and Alana Maurushat; New Zealand by David Vaile and Alana Maurushat; and Canada by Suzanne Palko and Alana Maurushat. The writing of these sections, however, for the purpose of this paper was done solely by Alana Maurushat. Any errors and omissions are those of Alana Maurushat.

** Jan Trzaskowski, Professor, Copenhagen Business School, author contact: jt.jur@cbs.dk.

*** Hadeel Al-Alosi, Lecturer, Western Sydney University, author contact: H.Al-Alosi@westernsydney.edu.au.
Open government data is exciting for many reasons. First, it allows for transparency and accountability in ways that extend beyond mere Freedom of Information requests for data. Because open government data is premised on creative commons licenses it allows anyone anywhere the ability to view information, to run analytics, to share information, and to produce innovative products based on the underlying data. The opening of datasets in machine readable linked data is of particular importance to university and private industry researchers as it opens hundreds of thousands of previously private datasets to be used for new research. Moreover, the more advanced jurisdictions have provided portals, tools, information, explaining how to best use the data. You don’t need to be a data scientist or highly trained statistician to run analytics on these datasets. Results are often turned into visualisations that are easier to comprehend than mere datasets. The goal then is to turn data into information and information into insights that allow for evidence base policies.

This paper summarises the results of research into the legislative, policy, regulatory and operational enablers utilised in selected advanced jurisdictions (UK, USA, France, Canada, New Zealand, Denmark, and the Netherlands) as identified in the Open Data Barometer and similar indexes. Australia is not considered in this article as at the time of review of the literature and data, it was not considered a global leader in the field. It is hoped that this paper will promote Open Data, and offer a view on how to encourage a culture of data sharing that can help inform future strategic developments in other jurisdictions. An ancillary goal is to encourage governments to continue to not only open government datasets but to do so in ways best compatible for researchers and industry to fully capitalise on the data to develop new products and innovative services as well as to heighten evidence base policy.

Communications were made with government agencies, open data departments and organisations in these jurisdictions in the period of December 2016 to the end of February 2017 to seek direct input as to how the frameworks have operated in practice. We contacted
many entities in the United Kingdom\textsuperscript{1}, United States\textsuperscript{2}, France\textsuperscript{3}, Canada\textsuperscript{4}, and New Zealand\textsuperscript{5}. Organisations did not wish to be identified with their specific comments. As such, insights gained from these communications is embedded within the analysis but is otherwise unattributed.\textsuperscript{6}

**Background**

Open data and open government data are important not only for general evidence-based policy but also for specific fields such as law and criminology. Often corrections and prison data are released to the public in summarised formats. Increasingly, agencies are releasing data in both summarised and raw/unaggregated data formats. The unaggregated data formats allow for increased and more sophisticated analytical use. In the context of data related to law and justice, justice data often includes corrections data, courts and sentencing data, law enforcement data, data specific to indigenous people, and victimization data. Within these types of data, there may be hundreds of different datasets which can be combined for analysis. For example, the US Bureau of Justice released which is known as ‘Law Enforcement Agency identifiers Crosswalk.’ The Crosswalk data:

\begin{quote}
provide[s] geographic and other identification information for each record included in either the Federal Bureau of Investigation's Uniform Crime Reports (UCR) files or BJS's Directory of Law Enforcement Agencies. The variables contained make it possible for researchers to take police agency-level data, combine them with Bureau of the Census and BJS data, and perform place-level, jurisdiction-level, and government-level analyses.\textsuperscript{7}
\end{quote}

\textsuperscript{1} United Kingdom: Leeds Council, Data Mill North, Scottish Cities Alliance, the City of London, the Open Data Institute, Socrata, the Office of the Prime Minister's Cabinet, and the Department for Environment, Food and Rural Affairs.

\textsuperscript{2} United States: The Obama Administration, the city of San Francisco, GovDelivery and the Policy Lab.

\textsuperscript{3} France: ETALAB, OPENDATA France, Data Gouvernance France, the French Information Industry of Online Information, General Secretary for Modernisation of Public Action, Marie de Paris, and Atelier Parisien d'Urbanisme.

\textsuperscript{4} Canada: Treasury Board Canada, Treasury Board of Ontario, and the city of Toronto.

\textsuperscript{5} New Zealand: The Department of Land Information New Zealand (LINZ), Department of Internal Affairs, Office of the Government Chief Information Officer, State Services Commission, the Ombudsman Office, New Zealand Data Futures and universities.

\textsuperscript{6} Communications data on file with author.

\textsuperscript{7} See Government of the United States of America, Bureau of Justice Statistics <https://www.bjs.gov./rawdata.cfm#law>
In a recent study in factors of judicial decision making, researchers examined open data from over a thousand cases by eight judges, combined with sentencing data, and other seemingly extraneous information.\(^8\) For example, the study found that prisoners were more likely to be granted early parole or a more lenient sentence at the start of the day or immediately after a break in court proceedings such as lunch or coffee. The researchers studied other factors such as severity of the prisoner’s crime, prison time, sex and ethnicity which was found to not affect the rulings in the same way as more innocuous factors. The study can now be modified to use data analytic tools to study hundreds of thousands of cases in varying jurisdictions to see if these factors hold true in other jurisdictions, and whether such factors equally influence civil proceedings.

Another example of evidence-based law looks at the European Union’s decision to mandate that clinical study reports for medical products are openly available for researchers.\(^9\) Normally, only abbreviated medical journal articles are published; the EU now publishes the full studies. While not a perfect solution to the problem with pharmaceutical commercial sponsorship of medical research, there is more transparency. Open access to such studies can lead to information for product liability, and potentially open up avenues for public redress where medical products have led to undesirable outcomes. These are merely two examples of hundreds of new ways that law can benefit from opened government data. For other countries, open government data means free and unfettered access to legal databases that include court decisions and legislation such as what is found on Austlii and Canlii. Other jurisdictions (many in Europe and the United States) still do not have free and open access to legal documents.

What is open data exactly? There are standardised definitions which are listed below.

Open data\(^10\) is data that can be used, shared and built-on by anyone, anywhere, for any purpose.

Open Government Data\(^11\) is:

- Data produced or commissioned by government or government controlled entities

---

\(^8\) Shai Danziger

\(^9\) Peter Doshi and Tom Jefferson, ‘Clinical study reports of randomised controlled trials: an exploratory review of previously confidential industry report’ (2013) 3(2) BMJOpen 1.


• Data which is open as defined in the Open Definition\textsuperscript{12} – that is, it can be freely used, modified, and shared by anyone for any purpose (subject to requirements that preserve provenance and openness) (Open Government Data at https://opengovernmentdata.org)

These definitions are important in that often there are misconceptions about information sharing and open data. For example, if an agency does not allow the data to be used in a commercial application, or if an agency is charging for the data, this is not open data. Some data for reasons of privacy, security, and commercial sensitivity has restrictions on if and how it is shared. However integral to this article is a recognition that not all data constitutes open data from creation and to progress the open data agenda enablers that facilitate the transformation of data to open data are required. This approach is recognised in the enablers identified in all leading jurisdictions.

Government may share data in three ways: 1. Internally within an Agency, (one department or a cluster of departments with similar sharing restrictions) 2. Agency to Agency (between government departments or authorised third party), or 3. Agency to public (between government agencies to the public). Under the definition of open data, data is only truly open when it is available to be accessed, used and shared in all of the above ways. An additional aspect that is increasing in importance is the use of external non-government organisations to provide services on behalf of government. In these cases there may also need to be appropriate legislative and policy frameworks supporting the flow of information from those providers to agencies.

**Selected Indexes and Countries**

There are many projects and indexes looking at open data but not all are directly relevant to this article.\textsuperscript{13} This article considers six different measures but relies primarily on the work of three different measurements of the extent to which jurisdictions have implemented Open Data. These are: *Open Data Barometer (2015)*, *Global Open Data Index (2015)*, and *OECD*.


\textsuperscript{13} Open Data 500 Global Network and the Govlab Index on Open Data study and compare companies’ use of Open Data and track open data companies with the goal to “improve people’s lives by changing how we govern, using technology-enabled solutions and a collaborative, networked approach”. The *World Justice Project Open Government Index* measures government openness based on publicized laws and government data, right to information, civic participation and complaint mechanisms. The scores and ranking draw on 78 variables derived from over 100,000 surveys and expert questionnaires for each country.
OURdata Index on Open Data (2014), and the Global Right to Information Rating (2015). The study further considered whether countries were members of the International Open Data Charter, the World Justice Open Government Index and the G8 Open data Charter (2013). By examining the leading jurisdictions for these rankings, countries were selected to examine with additional weighting given to the Open Data Barometer as it is the only indicator that assesses impact. Those countries were: The UK, France, the US, Canada, New Zealand, the Netherland and Denmark.

**Open Data Barometer (ODB)**\(^{14}\) is an expert assessment system that is scored by peer-reviewed local expert survey, a government self-assessment via a simplified survey and secondary data selected to complement the surveys to assess ‘Readiness’ portion of the assessment (data from the World Economic Forum, World Bank, United National e-Government Survey and Freedom House). Open Data initiatives are assessed by:

- **Readiness**: How prepared are governments for open data initiatives? What policies are in place?
- **Implementation**: Are governments putting their commitments into practice?
- **Impact**: Is Open Data being used in ways that bring practical benefit?

ODB is assessed across fifteen types of datasets: map data, land ownership, national statistics, detailed budget, government spend, company register, legislation, public transport timetables, international trade, health sector performance, primary or secondary education performance, crime statistics, national environment statistics, and national election results. ODB is the only study that assesses impact.

**Global Open Data Index (GODI)**\(^{15}\) is a crowd-sourced indicator of the openness of government datasets where information is gathered through the Open Data Census. The index is produced by the Open Knowledge Foundation and relies on contributions from civil society members and open data practitioners globally (through non-probability sampling technique – ‘snowball sample’). Any member of the public may contribute to the index which is later peer-reviewed and checked by a team of expert country editors, and lastly there is a public review.

\(^{14}\) Open Data Barometer (ODB), *The Open Data Barometer* <http://opendatabarometer.org/?_year=2016&indicator=ODB>.

The Index relies on the assessments of ten types of datasets: government budget, company registers, election results, emissions of (air) pollutants, legislation, national map, postcodes, government spending, national statistics, and transport tables.

**OECD OURdata Index on Open Data (OECD OGD)**\(^{16}\) is an indicator produced by the OECD that uses both an ex post and ex ante analytical framework for OGD initiatives around a related set of data in order to map initiatives across OECD countries. The common set of metrics can then be applied to assess the impact and value created from Open Data. Open Data is analysed in three critical areas – openness, usefulness and re-usability.

The index includes analysis of nine types of datasets: business information, registers, patent and trademark information, public tender databases, geographic information, legal information, meteorological information, social data and transport information.

**International Open Data Charter**\(^{17}\) was established in 2015 and builds on the G8 Open Data Charter, signed by G8 leaders in July 2013. The Charter is a collaboration between governments and data experts, and is underpinned by six principles to improve the access, release and use of data:

- Open by default
- Timely and comprehensive
- Accessible and usable
- Comparable and interoperable
- For improved governance and citizen engagement
- For inclusive development and innovation

**World Justice Open Government Index (WJ Open Government Index)**\(^{18}\) is an indicator of government openness based on four dimensions: publicized laws and government data, the right to information, civil participation and complaint mechanisms. The scores and rankings come from household surveys (over 100,000) as well as in-country expert questionnaires.

---


The index provides the “perspectives of ordinary people as they interact with their governments.”

The Global Right to Information Ratings (GIIR)\(^{19}\) is a programme which comparatively assesses the strength of legal frameworks for the right to information from around the world which is based on 61 indicators. The rating measures the legal framework based on clusters of indicators: Right of Access, Scop, Requesting Procedures, Exceptions and Refusals, Appeals, Sanctions and Protections, and Promotional Measures. A pilot application was conducted to test the framework, as well as looking at international standards and comparing them to countries right of information laws. Many of the local experts have a background in journalism and/or privacy. The ratings measure the legal frameworks; they do not measure their implementation, how they function in practice, or their impact.

G8 Open Data Charter\(^{20}\) was signed by the G8 leaders on 18 June 2013. The Open Data Charter sets out 5 strategic principles that all G8 members will act on. These include an expectation that all government data will be published openly by default, alongside principles to increase the quality, quantity and re-use of the data that is released. G8 members have also identified 14 high-value areas – from education to transport, and from health to crime and justice – from which they will release data.

The findings and rankings from these studies are provided below in Table A.

**Table A: Global Ranking Comparisons of Leading Jurisdictions**


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Yes</td>
<td>8</td>
<td>34</td>
<td>Yes</td>
</tr>
<tr>
<td>United States</td>
<td>2</td>
<td>8</td>
<td>9</td>
<td>No</td>
<td>11</td>
<td>57</td>
<td>Yes</td>
</tr>
<tr>
<td>France</td>
<td>2</td>
<td>10</td>
<td>2</td>
<td>Yes</td>
<td>17</td>
<td>95</td>
<td>Yes</td>
</tr>
<tr>
<td>Canada</td>
<td>4</td>
<td>17</td>
<td>5</td>
<td>No</td>
<td>7</td>
<td>49</td>
<td>Yes</td>
</tr>
<tr>
<td>(City of Edmonton - Yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>5</td>
<td>3</td>
<td>19</td>
<td>No</td>
<td>4</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>6</td>
<td>NA</td>
<td>15</td>
<td>No</td>
<td>2</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>6</td>
<td>8</td>
<td>25</td>
<td>No</td>
<td>5</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

France is noticeable in its rank change from 2013 to 2015 in the ODB (up 8 ranks) and is the only other country surveyed to also adopt the International Open Data Charter; they have adopted the G8 Open Data Charter as well.

Canada has adopted the G8 Open Data Charter and the City of Edmonton, Alberta, has adopted the International Open Data Charter.

This table highlights the leadership of the United Kingdom. It is the only country to score in the top 5 across these first three Open Data measurements. They were also one of the original adopters of the International Open Data Charter 2015\(^{21}\) and the G8 Open Data

Charter 2013. The United Kingdom is also the only country to Score 100 in Readiness, Implementation and Impact (ODB). As a result, this article focuses more heavily on the UK than other jurisdictions.

Each jurisdiction’s open data policies are organised by legislation, responsibility and coordination, policies, regulatory settings, operating environment and a selection of examples of open data projects, outcomes and applications.

**United Kingdom**

**Legislation**


The DPA provides a data protection regime that governs the collection, use and disclosure of personal data. The DPA covers ‘personal data’ (PD) via data protection principles, and deprecates retention of PD for longer than absolutely necessary, or release of PD unless collected under narrow circumstances, requiring redaction of records or removal or editing of information to reliably prevent identification of individuals where the source derives from personal data. Sensitive personal information has higher protection, and transfer outside the EEA is not permissible without adequate protection. The ICO can under the DPA prosecute offences, conduct audits, make orders and report to Parliament. The main intent is to protect individuals against misuse or abuse of information about them through encouraging best data management practices.

The FOI Act provides four main objectives: openness and transparency; accountability; better decision making; and public involvement in decision making. The FOI Act and its s45 Codes of Practice cover ‘disclosure of information held by public authorities or by persons providing services for them’, creating a potential for citizen enforcement of the right to

---


access certain information. The Protection of Freedoms Act 2012\(^\text{25}\) amended the FOIA to create a ‘right to data’ comprising new duties for certain public authorities to provide datasets of factual management information in a re-useable form and with a licence permitting re-use, in response to requests, and to continue to publish them. Re-useable means machine readable based on open standards.

The PRA\(^\text{26}\) operates in conjunction with the FOI Act to make arrangements for the selection and transfer of public records to the National Archive. The FOI Act determines when public records are released for public access.

The Re-use Regulations\(^\text{27}\) creates a specific OD driver requiring information be made available for reuse in machine readable format using open data standards and by default, the Open Government License

**Responsibility and Cooperation**

In the UK the ultimate responsibility for achieving open government (and within that open data) rests with the centralised authority of UK Cabinet Office, the work of which is largely coordinated through the Chief Data Officer\(^\text{28}\) who has the role of championing open data, driving the use of data for government decision making and by setting standards and principles for open data including the enforcement of set standards.

Open data initiatives are supported by a culture of openness at all levels of government (top, middle and bottom) through centralised, regionalised and localised efforts.

The accountability and transparency mandates in the Prime Ministerial Letters of 2010-2012 required OD take up by both centralised Departments and local government. There is also now vigorous regional activity with council clustering including Leeds and the North (Eg. Data Mill North), Bristol, London, and in 7 Scottish cities (Eg. Scottish Cities Alliance). The


role of harnessing and championing data at the localised level is folded into the role of Chief Information Officer or Chief Officer ICT\textsuperscript{29} (Eg. Leeds)

There are dedicated open data chief officers at the national, regional and local levels. The Information Commissioner Office\textsuperscript{30} is responsible for the oversight and integrity of open data through the establishment of the Open Government License, best practices, complaint handling, and determinative powers together with national and international leadership, cooperation and advocacy.

**Policy**

Prime Minister David Cameron issued Ministerial Letters (2010-2012)\textsuperscript{31} to every government department calling for greater transparency through specific commitments to making both information and datasets on government spending, procurements, crime data and more – all openly available to the public.

Government commitments or obligations at a national level include those under the Open Government National Action Plans,\textsuperscript{32} represent a policy foundation for Open Data putting release of public sector data into an accountability and transparency context. The 2012 Cabinet Office White Paper Open Data: unleashing the potential\textsuperscript{33} set out how the Coalition government aimed to ‘put data and transparency at the heart of government’.

The *Public Data Principles* remain another key policy document (Public Sector Transparency Board: Public Data Principles).\textsuperscript{34} There are 14 Principles which guide open data. These are important as other jurisdictions have Charters or similar documents that do not go as far as the UK. The selected principles are ones not readily found in other jurisdictions, but which contribute to the UK’s success in open data in readiness, implementation and impact:


\textsuperscript{30} Information Commissioner’s Office, <https://ico.org.uk>.


• Public data policy and practice will be clearly drive by the public and businesses that want and use the data, including what data is released when and in what form.

• Public data will be published using open standards, and following relevant recommendations of the World Wide Web Consortium (W3C)

• Release data quickly, and then work to make sure that it is available in open standard formats, including linked data forms

Commitments at Cabinet level were in response to the Prime Ministerial Letters of 2010-2012 which mandated Department Ministers to release a list of specified data sets for the purpose of accountability, and transparency of public spending and related decisions. At the subnational and council level, the Local Government Transparency Code 2015 mandates local authorities make plans and publish a number of open datasets, including data on spending (£500 and above) and contracts, senior salaries, grants to voluntary, community and social enterprise sector, public land and property assets, and other information, and report to the Department for Communities and Local Government (DCLG) on progress on the plans.

A main policy objective was to require adoption of ‘open by default’ for these data sets. The coverage is mainly ‘government to public’, in keeping with the accountability motive, with ‘machine to machine’ APIs increasingly expected. There appears to be little emphasis on ‘Agency to Agency’ release in UK Open Data literature. This may be due fewer restrictions on information sharing between agencies and programs to encourage information sharing. For example, many UK agencies participate in multi-agency working and information sharing (MASH). The same over-arching legislative and policy frameworks such as FOI, and Data Protection apply to national, regional and local agencies creating less information sharing limits between agencies, and between national and sub-national agencies. In the UK, data is open in its truest sense: for internal agency use, agency to agency, and agency to public.

Regulatory Settings


Responsibility for monitoring, enforcement and assessment varies according to the agency and level. 2012’s Open Data White Paper committed Cabinet Office to quarterly Written Ministerial Statements on progress against the Public Data Principles as set out by the Public Sector Transparency Board. Departments publish Open Data Strategies setting out their programmes of work towards embedding Public Data Principles as business as usual, and commitments on publication of a number of datasets. Progress made in these strategies as well as the departmental commitments set out in the Prime Minister’s letters of May 2010 and July 2011 on transparency and open data were also reported on as part of the evidence of progress.

FOI Act and its s45 Codes of Practice create a practical framework for right of access to information. A refusal or delay to provide information to the citizen triggers significant effort at the data host, and ICO adjudication. There is an incentive to avoid such case-by-case FOI effort by publishing the whole data set pre-emptively, if the capability exists and the data is suitable. Communications with open data experts in England confirms that this Open Data release path can reduce case-by-case FOI burdens, motivating an Open Data culture.

Operating Environment

Open data was initially driven by civil society’s requests for specified datasets and by Ministerial Letters (2010-2012) from David Cameron to every government departments calling for greater transparency through a specific commitments to making both information and datasets on government spending, procurements, crime data and more, openly available to the public.

Business and entrepreneurial activity is however not directly linked to Open Data release for accountability and transparency; there is generally more scope for widely valued and commercial uses of Open Data where the data sets covers issues of more general and routine

---


interest, like garbage collection nights (Data Mill North)\textsuperscript{40}, travel times (TfL)\textsuperscript{41} and food and agricultural markets (DEFRA)\textsuperscript{42}. The Open Data model has moved beyond internal accountability, compliance and releases to a more supported and collaborative OD culture.

In 2014 the UK authorised 1.5 million pounds to fund open data related projects. Since then there has generally been limited injections of extra funds to publish data sets as Open Data, with exceptions where central agencies have wanted to ensure a certain data set was available urgently. This ‘Business as usual’ model raises questions where there are continuing costs of conversion of legacy data sets and no resources for redevelopment of software to an ‘Open by Default’ model, especially where there are limited local practical benefits for the unfunded efforts to release sets or redevelop systems.

According to our engagement with agencies measurement has been done by quantity of datasets released with anecdotal evidence of cost reduction. There was a general sense that agencies did not have sufficient funds to do impact measurements.

There appear to be few if any binding quality requirements (apart from reliable exclusion of personal data, security sensitive information and some commercially confidential material). Raw ‘data exhaust’ datasets with known flaws are sometimes permissible to release. Quality issues were reported to be increasingly dealt with by expecting high quality data set metadata descriptions, so machine and human know what to expect.

Testing and Evaluation of Open Data sets must all be done under the standard project procedure before release. The publisher must assess compliance with the Data Protection Act, and take whatever action is necessary. This appears to include identifying those data sets that are not suitable for Open Data release for privacy or related reasons, and those where anonymisation/de-identification may fully address the DPA issues.

There many useful documents on Anonymisation and De-identification. These include the Information Commissioner Office’s ‘Anonymisation: Managing Data Protection Risk 2012,’

\textsuperscript{40} Data Mill North, <https://datamillnorth.org>.


and more recently the UKAN collaboration production of Anonymisation Decision-Making Framework 2016.43

The Open Data Champions in the organisation are responsible for testing and evaluation. The data owners may not be in a position themselves to make these assessments or do this evaluation.

Data Scientists and other key technical people are in short supply, especially in the regions, and there are as yet few tertiary courses for the multiple skill sets required. The rare experts that are employed often occupy multiple roles in public and private sectors, promoting policy, standards and resources in the public sphere and implementing project solutions in consultancy roles. However, training is offered to data custodians within the centralised Departments more generally; this focussed on privacy, anonymization of personal information, and differences between FOI and Open Data.

Hubs of clusters of expertise appear to be important to UK’s success, including around Leeds and the North, Bristol, London, and the Scottish 7 cities. Leeds for instance started the Leeds Data Mill, but rebranded this as Data Mill North when it became recognised as a functional hub of Open Data publication capacity and expertise for the region. Networks of Open Data champions and civil servant organisations also play a role in culture and knowledge sharing.

The UK provides a centralised open portal, Data.gov.uk44 platform. Licensing in UK tilts strongly to use of the Open Government Licence (OGL),45 especially centrally. Traditional ‘Crown copyright’ and Creative Commons CC-BY type licences were not considered appropriate for public sector Open Data release; OGL draws from these but is tailored to such releases. It also simplifies analysis of licencing implications, because it is the common default (Canada has modelled its license on the OGL).

Detailed and standardised metadata (Dublin Core Metadata Standard/ISP Standard 15836-200946 or W3C47) was seen to have potential to help identify and deal with issues about data

quality and characteristics, by putting both human and machine ‘on notice’ about its specific quality and other characteristics including the limitations of the dataset.

Examples of Open Data Projects, Outcomes and Applications

OPENDefra\(^{48}\)

The Secretary of State for the Department for Environment, Food & Rural Affairs (Defra) set a challenge for the department to transition to a more open, collaborative and data-driven organisation resulting in OPENDefra. OPENDefra was a collaboration of internal and external participants (Eg. Open Data Institute) were able to realise the release of over 11,000 datasets in 18 month (8000 specific datasets were mandated to be opened by the Cabinet Office). In this way, the Cabinet mandated a quantitative dataset quotas to be opened in a specified time. The big catalyst to opening the data came from the realisation that the data had potential uses and engaged users outside of the Department.

The release of high quality data, for example the Light Detection and Ranging (LIDAR)\(^{49}\) data yielded interesting results. LIDAR is the Environment Agency’s 3D height data used by the agency for flood modelling. The dataset was mandated to be opened. When the dataset was opened and released it found its way into a variety of applications and experiments including resources for school, the game Minecraft, modelling of snowfall for scientists working on climate change, in urban planning and civil engineering to help plan and manage infrastructure by transport, energy and utility companies. Previously this data was a revenue generator causing some concerns over revenue reduction if the data were to be opened. However, while revenue disappeared they saved money by opening the data. Prior to open data, many of the flood predictions were done by companies using less reliable datasets. These models and applications in turn had to be carefully reviewed due to data quality issues. Use of the high quality LIDAR data has alleviated testing and evaluation costs.


OPENDefra also resulted in an appraised approach to privacy. The National Family Food Survey\(^{50}\) conducted a privacy impact assessment\(^{51}\) which was later been used as an example of a model approach. It was later published and there remains a version open to public comment to provide feedback for current and future use. This PIA is considered to be a model for future opening of datasets containing confidential personal information.

A privacy impact assessment was performed, published and there is a version open to public comment for the PIA to provide feedback for current and future use. The PIA is considered to be a model for future opening of datasets containing confidential personal information.

**Regional and Local Skills and Training Hub: Data Mill North**

Data Mill North publishers were created by Leeds City Council, and recently spun off and rebranded independently. Along with the Leeds Council and Innovation Lab, they are seen across the UK as the most innovative and progressive hubs of innovation. Now a regional publication hub and centre of expertise for the North England region. It needed a critical mass of staff, who are hard to support both in budget and retention, but which is impossible at the individual council level. It works over a large region in the North of England, joins various councils together, common resources, cross fertilisation, critical mass of different initiatives being worked up, coming on line, and going operational.

**Freedom of Information Requests: Leeds Council**

The Leeds Council had the benefit of reducing costs associated with Freedom of Information Requests (FOR). Leeds received 10 FOIs per month for business rates data, taking 30 hours per month, nearly a whole week for FOI on this data type alone. By publishing it online first, over time the recipients got used to it and making less FOI requests. Where an FOI request was made the Council could simply direct them to where the information/data was openly available online. This service now also publishes additional data, guided by actual demand at the larger scale of online usage. This is all done with fewer resources than before. In the absence of other funding, this sort of outcome is very helpful. It relied on strictly enforceable

---


FOI rights - they could not be avoided by delay or refusal, so Open Data was a viable alternative.

**United States**

**Legislation**


The DATA Act\(^{52}\) requires annual federal spending to be reported as open data in machine-readable format. The Act was first passed in 2014 and is being implemented incrementally with the final implementation date of May 2017.

The Executive Order\(^{53}\) mandates open by default for new and modernised government information. It is important to note that the title of the EO includes the term “machine-readable”. This resonates with the US viewpoint that opening data is only the first step. Advancing a data-driven government and economy requires open data, machine-readable format, with the legislative framework and sufficient resources to utilise big data analytics.

Related acts include Freedom of Information Act\(^{54}\) which instructs agencies to “adopt a presumption in favour of disclosure” and “take affirmative steps to make information public.” The Open Government Act 2007\(^{55}\) that amended sections of the *Freedom of Information Act*.

In California Open Data is required under s. 6253.10 which inserted “open data” in to the California Public Records Act 1968.\(^{56}\) The Act requires any agency that voluntarily posts any public record described as “open data” to make that record available in an open format.


(i.e., retrievable, downloadable, indexable, searchable, platform independent, machine readable, free of charge, and retaining its compiled data structure and definition).

The city of San Francisco has created its own open data legislation including the San Francisco Sunshine Ordinance 1999\textsuperscript{57} and the Open Data Directive 2009.\textsuperscript{58}

**Responsibility and Cooperation**

President Obama issued the Open Data Executive Order in May 2013 along with The Open Data Policy\textsuperscript{59} released through the Office of Management and Budget\textsuperscript{60} and the Office of Science and Technology Policy.\textsuperscript{61} The real push for open data and machine-readable data came from civil society as was also the case in the UK and France.

The US established a governance framework with a list of roles and responsibilities. Like the UK and France, there is a Chief Information Officer\textsuperscript{62} (similar to the Chief Data Officer) who is responsible for policy and oversight.

The Office of Science and Technology Policy has a Chief Data Scientist.\textsuperscript{63} The OSTP is responsible for championing open data across designated fields of energy, education, finance, public safety and global development. There is concern under the Trump Administration that there will be budget cuts to OSTP as well as a potential change in policy potential in open government data as a whole, and certainly of opening certain datasets. Energy and environmental datasets appear to be at the most risk.

\textsuperscript{57} 67 Cal Code <http://administrative.sanfranciscocode.org/67/>.
\textsuperscript{61} Government of the United States of America, The White House, Office of Science and Technology Policy <https://www.whitehouse.gov/ostp>.
The Office of Government Information Services (OGIS)\(^{64}\) provides mediation service for disputes between FOIA requesters and government agencies. They also review and provide input into policies and procedures as they relate to agencies under the FOIA.

The United States is unique in that it promotes open data for democratic transparency and accountability reasons (found in all jurisdictions) as well as for economic reasons. Open Data is considered a key future economic driver. Open machine-readable data is considered an essential element for businesses, governments and researchers to make new products, services, create jobs, and establish new businesses. Open data enhances innovation policy.

Many US States and cities, such as the State of California and the city of San Francisco, also have Chief Data Officers or the equivalent that are supported by legislation, policy and a sufficient budget to drive data initiatives.

**Policy**

Sent on his first day in office, President Barack Obama sent a Memorandum\(^{65}\) for the Head of Executive Departments and Agencies on Transparency and Open Government. The Memo commits to establishing “an unprecedented level of openness in Government,” and argues that government should be transparent, participatory, and collaborative. Along with these three principles, this memo directs the Chief Technology Officer, Director of the Office of Management and Budget (OBM), and Administrator of General Services to develop recommendations for “an Open Government Directive,” and it instructs executive departments and agencies in the federal government to act accordingly.

The *Open Government Directive* 2009 was issued by the Obama Administration and directs executive departments and agencies to take specific actions and to implement the principles of transparency, participation and collaboration which including publishing government information online in open formats.

---


The Open Government Partnership: National Action Plan for the United States of America66 was first published in 2011 and provides a national action plan coupled with self-assessments and status reports. It lists new initiatives that include a platform for the public to directly petition the White House (https://petitions.whitehouse.gov/),67 reform of FOIA, streamlining declassification, and disclosing regulatory compliance information. The Action Plans “to manage public resources more effectively,” initiatives are planned to increasing transparency around extractive industries, federal spending, and foreign assistance. And, “to improve public services,” initiatives build on Data.gov and other websites for communication and collaboration.

Subsequent plans and reports have been published under the Open Government Partnership, particularly the 2013 Self-Assessment Report.68 These plans include support for subnational and global development in this space (similar to Canada’s Action Plan).

The Digital Government Strategy69 compliments open government initiatives by setting goals and timelines to modernise the delivery of digital services. It focuses on information or data (accessible through web APIs), shared platforms (including open content management systems and enterprise wide asset management & procurement), internal and external customer service (e.g., improving .gov domain and mobile access), and security (including personal privacy and mobile security requirements).

It is unknown what direction Open Data will take with the Trump Administration. As of February, 2017 there have been a number of key datasets removed from the main Federal portal data.gov. The removals have been tracked by the Sunlight Foundation.70 Replicas of the datasets were made and are stored by the Internet Archive, numerous libraries and universities and organisations outside of the United States. The datasets are still openly

available, but not through the official data.gov portal. There have been no legislative amendments, regulations or government policies announced in relation to open data as of yet.

The city of San Francisco is exemplary in its adoption of law, policy and support for open data. The Open Data Policy 2010 (Administrative Code Chapter 22D, Open Data Policy)\(^{71}\) codifies in law the policy requirements from the mayor’s previous Open Data Directive. It requires departments to “make reasonable efforts” to make datasets available through the city’s portal and regularly review their progress in doing so. The city’s Committee on Information Technology (COIT) is to establish rules and standards for public disclosure (within 60 days), and “balance the benefits of open data… with the need to protect from disclosure information that is proprietary, confidential, or protected by law or contract.” The city operates the DataSF portal.

**Regulatory Settings**

Like the UK and France, the role of the Deputy Chief Technology Officer for Data Policy, and the Chief Scientists in the OSTP are considered pivotal. The role of CDOs in nations share many common goals: help shape policies and practices, enable a culture, and assist in developing guidelines, standards and licenses. In the US, however, the role of CDO equivalent is also to “maximise the nation’s return on its investment in data” and to “help recruit and retain the best minds in data science to join in serving the public.”

Both California and the city of San Francisco have dedicated Chief Data Officers. For San Francisco the duties of the position are listed to include drafting rules and standards; prioritizing data for publication; coordination and maintenance of DataSF; and assistance to city departments, among related functions.

**Operating Environment**

While open data polices may produce long-term savings, implementation at the federal level is not cheap. In 2016, more than $80 million was requested for DATA Act implementation; $10 million was requested for pilot programs in the Department of Health and Human Services alone. Accurately or not, open data laws and policies are often seen by government agencies in terms of compliance and the additional costs of implementation.

\(^{71}\) [http://administrative.sanfranciscocode.org/22D/].
Implementation of open government data moved slowly at first in San Francisco, then improved significantly with legislated funded roles for this task.

Communications with open data organisations indicate that measurement is done through milestone achievements, and by both quantity and quality of apps/projects that have used the open data.

Agencies must incorporate privacy analyses into each stage of the information's life cycle. In particular, agencies must review the information collected or created for valid restrictions to release to determine whether it can be made publicly available.

As agencies consider whether or not information may be disclosed, they must also account for the "mosaic effect" of data aggregation. Agencies should note that the mosaic effect demands a risk-based analysis, often utilizing statistical methods whose parameters can change over time, depending on the nature of the information, the availability of other information, and the technology in place that could facilitate the process of identification. Because of the complexity of this analysis and the scope of data involved, agencies may choose to take advantage of entities in the Executive Branch that may have relevant expertise, including the staff of Data.gov.

While specific training programs were not mentioned by agencies, most noted that training of existing staff and hiring people with the requisite skill set were important enablers in creating an open data culture.

The United States central portal uses machine-learning to automatically consolidate datasets published on local and regional data portals and offering them on the national portal.

Data.gov and DataSF provide information on how to navigate the data but not provide the tools themselves. The information outlines how to search data, apply filters and recommended software languages and data formats are explained.

Examples of Open Data Projects, Outcomes and Applications

**DataSF (San Francisco)**

The US emphasises that data alone is not useful. Success depends on engagement. Many US efforts have included pilot studies and experimentation as well as published milestones,
progress reports and dashboards. The DataSF portal tracks the status of the dataset inventory, data plans and published datasets.

DATA USA

The most salient feature of open data is to publish the data as machine readable so that it may be linked to open data. The best example of this is DATA USA. Because the Federal Data.gov portal along with many State and local portals are machine-readable, issued in standard formats, and have clear licensing terms, the US private sector has greatly leveraged the open datasets. This is perhaps best illustrated through the recently launched DATA USA (April 4, 2016).

DATA USA is a free and open platform created collaboratively by MIT Media Lab, Deloitte and Datawheel (a Media Lab spinoff). The platform aggregates public data relevant to key issues providing what many consider to be the most comprehensive and easy-to-use open-source visualisation tool for public data. As one leading expert put it, “It’s essentially a one-stop shop for information that is easy to search, understand, embed, and build into new code.”

The author looked and experimented with many portals and applications, and did not find anything comparable to this system.

Code for America

Code for America is an excellent example for building capacity for regionalised and localised open government data. Code for America The Code for America is a foundation backed by the private and public sectors who “build open source technology and organise a network of people dedicated to making government services simple, effective, and easy to use.”

The Foundation selects 30 fellows each year to work with the Data Office in San Francisco to assist 10 American cities with open government data projects. Because these projects are open source, they are also made public so that anyone can contribute to the open source project, not just the 30 fellows steering the project. The project has inspired other global initiatives including Code for Africa. The Foundation has published an open data

---

73 DATA USA, <https://datausa.io>.
“playbook”\textsuperscript{25} based on its rich experience of working with sub-national authorities in the United States.

\textbf{France}

\textbf{Legislation}

As is the case with the UK and US, France has legislation mandating the release of open data (subject to limited exceptions).

The French legal framework consists of two main laws: LOI n° 2015-1779 du 28 décembre 2015 relative à la gratuité et aux modalités de la réutilisation des informations du secteur public' (\textit{Law on the free use and the modalities of the re-use of public sector information}),\textsuperscript{76} and Loi pour une République numérique 2016' (\textit{Law on the Digital Republic}).\textsuperscript{77}

The \textit{Law on the Digital Republic 2016} together with \textit{The Law on the Free use and the Modalities of the Re-use of Public Sector Information} mandates open by default; secure access to data for public researchers and statisticians; free access to data; machine-readable; free exchange of data between State administrations (prior some State administrations would have to pay a fee to access and use other State’s datasets); standard API; and that specific datasets by open which are the INSEE data, Public Administration data, energy related data, and legal jurisprudence data.

\textbf{Responsibility and Cooperation}

Open Government Data coordination function are centralised in France. \textit{Etalab}\textsuperscript{28} was established in 2011 where responsibility rests with the General Secretariat (for the Modernisation of Public Action) under the Prime Minister’s Office. \textit{ETALAB} coordinates the activities of public digital services, and its public institutions in order to facilitate the


\textsuperscript{76} (France) <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031701525&categorieLien=id>.

\textsuperscript{77} (France) <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000033202746&categorieLien=id>.

widest possible re-use of its public information. France has a State Chief Data Officer\(^79\) (2013) as does the greater city of Paris\(^80\) (2016).

**Policy**

The Atelier Parisien d’Urbanisme (APUR)\(^81\) made an internal decision to open datasets before the enactment of national law that made open data by default for France’s public administrations. As in the case of the United Kingdom, United States and Canada, there were publically made available Ministerial Letters (Vade-mecum)\(^82\) mandating the sharing and opening of data (2013). The legislative framework came after APUR’s move to open data, and the issuance of Ministerial Letters mandating open data.

France operates the open centralised portal, data.gouv.fr.\(^83\) Data is released under the License Ouverte,\(^84\) which as adapted from the Creative Commons License and is compatible with the UKs Open Government License and other Creative Commons licenses. Opening licensing law were the results of a French law mandating that public information be free and re-usable (Order No., 200-5-650). Currently this license has the unique restriction that data users may not deteriorate the content of the information or change meaning of words. France has announced that they will soon release 2 types of licenses for public agencies accompanied with detailed explanations of the licenses.

France released its National Action Plan: For a Transparent and Collaborative Government (2015-2017)\(^85\) along with the Open Government Partnership: Mid-Term Self-Assessment

---


**Regulatory Settings**

France is the first jurisdiction to have a State Chief Data Officer. France created a role of State Chief Data Officer whose mission is to organise a better circulation of data within public administration. The CDO is tasked with acquisition of essential data, contribute to and improve the quality of data, dissemination of tools, methods, guidelines to both comply with objectives of French law as well as to foster an open data culture. The CDO completes annual report on data governance.

**Operating Environment**

In France, the supporting guides along with the work of OpenData France, which is an association aiming at supporting municipalities involved in open data initiatives, have facilitated the enactment of an open data culture within agencies. OpenData France provides with a comprehensive set of supporting elements (advice, negotiation, development, tools for representation etc.). Data quality assurance is managed through disclosure of the possible inaccuracies and limitations of the dataset.

Etalab manages the portal www.data.gouv.fr⁸⁸ designed to collect and make available freely all public data.

The French use the Licence Ouverte⁸⁹ based on the Creative Commons licence. They are, however, currently developed two new types of licences for public sector agencies. In France there is a prioritisation of high-value datasets released with efforts made to consult with civil society and organisations. There is also the ability of a dataset and/or application to become certified. The certification is meant to instil confidence in users in the integrity and quality of the underlying dataset.

**Examples of Open Data Projects, Outcomes and Applications**

---


Opening Sensitive Datasets Through Sound De-identification and Anonymisation

France has historically had very strong privacy laws. In a field as sensitive as healthcare, the French decided to open healthcare. France introduced very recent legislation LOI n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé (Laws on the Modernisation of Medical System 2016) mandating that aggregated healthcare data be opened for researchers, and for analysis to improve the administration of the French medicare system. They are confident in their de-identification techniques, standards, support and training.

The government determined that healthcare data should be:

1) Opened up regardless of the potential use or re-use that could be made of it;
2) Opened regardless of the potential use or re-use that could be made of it;
3) Opened as granular as possible, while ensuring anonymity and complying with laws such as on a commercial confidentiality;
4) Made public whenever future surveys and research is funded by public means.

Canada

Legislation

The Access to Information Act 1983, the Federal Accountability Act 2006, Privacy Act 1985, and the Charter of Rights and Freedoms 1982 form the relevant legislative framework. As with the other leading jurisdictions, there is a combination of personal data protection coupled with FOI laws and accountability/transparency laws which contribute to the general open government and open data agendas.

Privacy frameworks are found at the national, provincial and municipal levels. Privacy law is not as strong (has not been updated to account for changes in technology) at the national level as applied to government entities when compared to the strong privacy and data protection

---

principles that apply to the provinces, and the federal framework that applies to the private sector. For example, the province of Ontario has a Municipal Freedom of Information and Protection of Privacy Act 2004⁹⁵ that is binding on all municipalities. The *Charter of Rights and Freedoms*, however, does speak to trust in the judiciary, accountability of politicians, right to a private life and a federal system of government which allows ideas to be tested at the provincial and territorial level, and then brought up to the federal level. In Canada, however, the real push for open government and open data is found in policy at the national, provincial and municipal levels.

**Responsibility and Cooperation**

The Chief Information Officer⁹⁶ of the Treasury Board of Canada⁹⁷ is responsible for delivering open government and open data in Canada together with the Departmental Information Managers from the various Federal government departments.

In the province of Ontario, this role is assumed by the Treasury Board Secretariat where there is a Deputy Minister for Open Government.⁹⁸

The Smart City of Toronto (in the province of Ontario) has its own policies, guidelines, portal and licenses.

**Policy**

Upon taking office, Prime Minster Trudeau issued an open letter⁹⁹ to Canadians expressing his commitment to an open and transparent government, one that is open by default. He also sent Ministerial Mandates¹⁰⁰ to each of his Cabinet members also containing language of openness, transparency and open by default.

---

In a similar move, Ontario Premier Kathleen Wynne wrote to each Cabinet Minister in the open Mandate Letters 2016\textsuperscript{101} outlining the key priorities for each ministry with some emphasis on the ‘digital transformation of government’. Open datasets are considered part of a larger picture of open government and specifically to better data collection, information sharing, evidence-based decision making and modernising public service delivery.

The Directive on Open Government (2014)\textsuperscript{102} is a non-binding document applicable to government departments. Although departments are not legally mandated to open government and open data, they act as a form of soft law that strongly influences action.

Canada released its Third Biennial Plan to the Open Government Partnership (2016-18)\textsuperscript{103} establishes 22 commitments based on four priority areas: 1. Open by Default, 2. Fiscal Transparency, 3. Innovation, Prosperity and Sustainable Development and 4. Engaging Canadians and the World. Each of the commitments sets milestones, information on how this is to be achieved and identified the lead department.

Datasets are released through the Open.canada.ca\textsuperscript{104} portal which operates in English and French. There are comprehensive guides, support, and instructions on APIs for computer programming languages and data formats. Canada adopted the UK’s Open Government License.

The province of Ontario has its own Open Data Directive\textsuperscript{105} that mandates that data should be open by default, high-value data should be prioritised for release, and that a data inventory be published online including a list of datasets which cannot be made accessible to the public along with a detailed explanation as to why this is the case. The Directive provides adopts the Open Government License and an Open Format Standard, and contains six Open Data Quality Principles which are a similar but less forceful version of the UK’s Data Principles. For example, in Ontario departments are encouraged to release data in a timely, coherent and interpretable manner but these remains concepts and are not as prescriptive such as the UK’s

and France’s requirements for “machine readable” or “release data quickly, and then adapt to open format”, and release in an international standardised format (W3C).

Ontario datasets are released through the Ontario Data Catalogue.\(^{106}\) The datasets are inventoried including those that are currently restricted.

Toronto’s OD policy is additionally aligned with Ontario’s requirements of Access by Design, and Privacy by Design. Executives are encouraged to release data and must release data where a formal Freedom of Information request has already been made or is in the process of being disclosed, as well as there the data/information has already been made available to the public.

**Regulatory Settings**

At the Federal level, the Treasury Board is responsible for monitoring and reporting on compliance with the *Directive on Open Government*.\(^{106}\)

Provincially, the role of monitor and report compliance lies with the Treasury Board Secretariat. The bulk of this role’s responsibilities are assumed by the Assistant Deputy to Minister for Open Government who is part of the Treasury Board Secretariat.

In Toronto, the Executives (Eg. City manager, City Clerk, and General Managers) are accountable for ensuring compliance with open data and open government policies. When an Executive determines that they are unable to comply with their role, they are meant to bring their non-compliance issues to the Open Government Committee for review and assistance.

**Operating Environment**

The 2016 national budget includes doubling the Treasury Board Secretariat’s budget for open government activities to deliver an ambitious open government strategy and to accelerate the provision of digital content, and to accelerate and make easier Canadians access to government data ($24.4 million over five years).

The implementation roadmap for Canada’s Open Government 2016-2018 Action Plan\(^{107}\) calls for the development of approaches for measuring open government performance by integrate performance indicators for openness and transparency into a Performance Management

---


Framework for Open Government. There is no indication that measurements of implementation or impact are currently done.

Open data is subject to both freedom of information and protection of privacy. Open datasets must be released with proper privacy mechanisms in place. The Information and Privacy Commissioners at the national and provincial levels provide guidelines around access to privacy, privacy by design including data anonymisation.

Canada has adopted the UK’s Open Government Licence model.

**Examples of Open Data Projects, Outcomes and Applications**

**Ministerial Letters or Equivalent at Federal, Provincial and Municipality Levels**

Open Government and Open Data leadership in Canada has been at the national and subnational levels. Ministerial letters and equivalent mandates were issued by the Prime Minister, Premier of Ontario, and Mayor of the city of Toronto. All levels of government encourage information sharing and open data by default.

**Province of Ontario**

The province of Ontario and city of Toronto embrace three data principles: Access by Design, Privacy by Design and Open by Default. These are considered complimentary principles, not competing principles. Ontario has an open government project tracker that allows the public to see the stage a project is at including planning, complete and implementing. Completed projects at this stage are largely policies and mandate letters, with crowdsourcing, data inventory and open government consultation underway. The category is complete as restricted and closed datasets are also listed.

**Denmark**

**Legislation**

The EU Public Sector Information Directive establishes a minimum set of rules governing the re-use and the practical means of facilitating re-use of existing documents held by public sector bodies of the Member States. The PSI Directive is implemented in Denmark through

---

Act no. 596 of 24 June 2005 as amended by Act No. 551 of 17 June 2008, and Act no. 553 of 2 June 2014. The Act neither mandates open data nor re-use; it merely suggests that public sector bodies may make documents available for re-use and provides principles governing such disclosure. It is also possible to apply for access to data by approaching the public sector body that administers the data collections (Section 4, subsections 3–6). Applications must normally be processed within 7 working days, and the applicant must be informed about possible significant cost involved with processing the application. Application must, if possible, be processed electronically (Section 5). If the application cannot be accommodated, the public sector body must inform the applicant about the owner of the data or from where license has been obtained.

_Danish Act on Processing of Personal Data 2000_ is administered by the Danish Data Protection Agency. The Act establishes relevant data protection rules, including the principle of proportionality, data minimisation and purpose limitation.

**Responsibility and Cooperation**

The Danish government is divided into three tiers: the state (‘staten’), five regions (‘regioner’), and 98 municipalities (‘kommuner’). The responsibilities for open data lie with each of the organisations. Open data is centralised, regionalised and localised as members from each of these cooperate as partners in the Open Data DK.

**Policy**

The Danish Government, Local Government Denmark (association of Danish municipalities) and Danish Regions (association of Danish regions) have in May 2016 entered into an agreement on a _Digital Strategy for 2016–2020_ (‘_A Stronger and More Secure Digital Denmark_’). The purpose of Open Data DK and the Strategy is to create transparent governance and support data-driven growth and productivity through by means of open and freely available data-platform with a view to publish their datasets.

Denmark has a tradition for national registries The Danish Agency for Digitisation under the Ministry of Finance is responsible for the Basic Data Programme (‘Grunndata’), under which local and central government are working to open registries.
There is no compulsory or recommended license or standard. There is no obligation to release the datasets free of charge.

**Regulatory Settings**

The Danish Business Authority is part of the Ministry of Industry, Business and Financial Affairs and the Agency for Digitisation is part of the Ministry of Finance. The Danish Agency for Digitisation which is also under the Ministry of Finance is responsible for the Basic Data Programme which have been mandated to open specific high value registries rolling out 2016 to 2018. These are the Civil Registration System, maps and other geographical data, the Central Business Register, and the Building and Dwelling Register. At a future date the expansion will include data on incomes, road infrastructure and financial statements of business.

**Operating Environment**

Open data for central, regional and local data is made available for the portal, Open Data DK (www.opendata.dk) and is built on the Open-Source software CKAN from the Open Knowledge Foundation.

Denmark has created a model where the centralised portal or data hub are financed through costs saved as a result of reduced costs for data hosting (previously not centralised) and data purchase (now free).

**Example of Open Data Projects, Outcomes and Applications**

A study undertaken by the Danish Government assessed direct financial benefits from opening utilities, address data, the Land Registry and the Central Business Registry and found that it cost two million EUR to open the data, but that the direct financial benefits from 2005-2009 were 62 million EUR.¹¹⁰

---

¹⁰ Open Data Denmark, [http://www.opendata.dk].
¹¹ Juliet McMurren, Stefaan Verhulst, and Andrew Young, *Denmark’s Open Address Data Set: Consolidating and Freeing up Address Data* (January 2016) available at [http://odimpact.org/static/files/case-study].
New Zealand

Legislation

Open Data is not required by law or by default in New Zealand. There is however a statutory basis in the Official Information Act 1982\textsuperscript{111} for requesting access to declared categories of information creates a principle of availability where, the Ombudsman may review complaints, and appeal enforcement. The view of the agencies consulted is that the legislation does not create a right to information or a legislative mandate to release data by default. It does, however, create a sympathetic environment to open data and data by default.

The Public Records Act 2005\textsuperscript{112} and others require creation and retention of certain information, and the Privacy Act 1993\textsuperscript{113} requires exclusion or protection of personal information.

Responsibility and Cooperation

The Open Government Data Chief Executives Governance Group reports the aggregate plans annually to the Ministerial Committee on Government ICT. Cabinet also invited the Minister of Local Government to write to local authorities and Local Government New Zealand informing the local government sector of the these decisions and encouraging councils, where they consider it appropriate, to take a similar approach. The framework home page includes Reports on the adoption of the framework each year from 2012-2015.

The CEO of Land Information NZ (LINZ) became the Government Chief Information Officer, with LINZ hosting the key Open Data program.

There are advocates for open data policies and legislation outside government, although the network is less well developed and extensive. Civil society and non-government organisations like the Open Data Catalogue\textsuperscript{114} the Open Data NZ Meetup\textsuperscript{115} and Open Government Ninjas\textsuperscript{116} promote open data at the national and local levels, and there is some participation in government supported entities like the New Zealand Data Futures Forum, now the Data Futures Partnership.

\textsuperscript{115}Open Data NZ Meetup, <https://www.meetup.com/Open-Data-NZ/>.
The Department of Internal Affair (DIA) also plays a key role by hosting the platform data.govt.nz. The State Services Commission (SSC), as the central agency overseeing the government sector, offers guidance on how to disclose data. The Ombudsman, amongst many other things, hears complaints by individual requesters against declined requests for release of information.

**Policy**

In the absence of specific Open Data legislation, a series of related policy documents from 2011 form the framework of the current support for Open Data in New Zealand. These are The New Zealand Government Declaration of Open and Transparent Government, and The New Zealand Data and Information Management Principles.

In 2011 Cabinet issued a Declaration that directs all public service and non-public service departments including NZ Police, NZ Defence Force, Parliamentary Counsel Office, and NZ Security Intelligence Service “to commit to releasing high value public data actively for re-use, in accordance with the Declaration and Principles, and in accordance with the NZGOAL Review and Release process.” Cabinet also encouraged other State Services Agencies and invited State Sector Agencies to do the same. Cabinet also directed Chief Executives to submit their plans to actively release public data to portfolio Ministers for approval.

High value datasets are to data be publicly released according to the New Zealand Data and Information Management Principles, which include principles such as Trusted and Authoritative, and Well Managed.

The Chief Ombudsman’s major report in December 2015, ‘Not a game of hide and seek’, made recommendations on OIA practices to support open data including support openness.

---


and accessibility (Recommendation 7), formats should enable easy reuse (Recommendation 21), proactive release of data (Recommendation 22), and managing risks that may arise from release (Recommendation 22).

New Zealand is additionally using a crowdsource method of developing policy around personal information and data used known as the Social License. The Social License is a partnership between New Zealanders and the government where people can contribute their thoughts on the contents of data guidelines.

**Regulatory Settings**

The CEO of Land Information NZ (LINZ) played a key role, becoming the Government Chief Information Officer, with LINZ hosting the key Open Data program.

**Operating Environment**

There was significant central budget funding for the NZ Open Government Data and Information Programme. For instance $300,000 extra provided for the two financial years to mid-2016, mostly for lead agency Land Information New Zealand (LINZ). It also went towards maintaining NZGOAL, and providing support to Creative Commons Aotearoa New Zealand.

NZGOAL specifies a review and release process which involves six Stages prior to Release: evaluation of copyright-related rights [101–109], evaluation of restrictions [112–118], selection of re-use rights [119–124], application of Creative Commons licence, or ‘no-known-rights’ statement [125–141], a moral rights check [142–144], and selection of formats, preferring non-proprietary ones [145–148].

NZGOAL requires ‘Anonymisation’ of personal information as well as an interrogation of the anonymization processes to ensure that they are rigorous.

---

122 Creative Commons Aotearoa New Zealand, <http://creativecommons.org.nz>.
Assistance on Anonymisation of datasets is the responsibility of the Office of the Privacy Commission.

This NZGOAL *(New Zealand Government Open Access and Licensing) Framework* (‘the Framework’) supports government agencies wanting to enable appropriate re-use of the agency’s material by licensing their copyright works or releasing non-copyright material for re-use. It includes a Software Extension based on the GPL and the MIT licence. New Zealand has not followed the UK path of a special Open Government Licence, instead choosing to adopt the most widely used Open Content licence, Creative Commons (CC-BY-4)

There isn’t a single catalogue or repository for open data in NZ. There are, however, many different platforms including: places to start looking’: GovHack,\(^\text{125}\) data.govt.nz,\(^\text{126}\) and the Open Data Catalogue.\(^\text{127}\)

**Example of Open Data Project, Outcome and Application**

The Social Licence is a unique experiment that provides meaningful, inclusive and ongoing engagement on open data. New Zealand provided a ‘Our Data, Our Way’ survey site. This tool seeks anonymous web user responses about benefit and trust levels for three brief scenarios. Each scenario has three increasingly privacy-intrusive proposed uses to which the user responds. The survey seeks feedback about trust factors and benefits of open data. This is a potentially meaningful, inclusive and ongoing engagement about open data.

**The Netherlands**

**Legislation**

What is particularly unique about the Netherlands is that access to government information is considered to be a fundamental democratic right, which is explicitly protected under Article 110 of the Constitution of the Kingdom of the Netherlands. By law, the government is required to share public information with the public upon citizens making a request. While the government may publish information on its own accord, the Dutch Government publishes considerably less data than the United States and the United Kingdom.


Under the *Government Information (Public Access) Act 1991*\(^{128}\), any member of the public is entitled to request information relating to an administrative matter. If the information is located with a public body or private company that provides services to a public entity, the authorities must respond within two weeks of the request. There are of course exemptions that allow the Dutch Government to withhold information it relates to international relations of the state, the "economic or financial interest of the state”, investigation of criminal offenses, inspections by public authorities, or personal privacy. These exemptions must be balanced against the importance of disclosure. Requesters who have been denied access can appeal to an administrative court that will then make the final decision.

**Responsibility and Cooperation**

Public access to government information has been a general legal principle in the Netherlands since 1980s. As mentioned above, it is the responsibility of the Dutch Government to make facilitate access to information. It has actively encouraged open data through policy and collaboration.

Within the Government, several ministries have taken the responsibility of providing open data, including the Ministry of Infrastructure and the Environment, the Minister for Foreign Trade and Development Cooperation has released data on international development, and the Ministry of Finance.

**Policy**

To promote open data, the Dutch Government has introduced several action plans. The most recent is the *Open Government in the Netherlands Action Plan 2016-2017*\(^{129}\). The aim of the Action Plan is to encourage “openness” and “transparency” within the public sector. There are a number of initiatives set out in the Action Plan, including the *National Open Data Agenda*\(^{130}\), which outlines the frameworks governing how data is to be made accessible & the quality requirements it must meet.


\(^{130}\) *National Open Data Agenda*<https://www.rijksoverheid.nl/documenten/kamerstukken/2015/11/30/kamerbrief-over-nationale-open-data-agenda-2016-noda>
Regulatory Settings

There is no one regulatory body responsible for monitoring open data initiatives in the Netherlands. It varies, depending on the project and are often in collaboration with several agencies. For example, the Ministry of the Interior and Kingdom relations, in association with all other ministries, is responsible for the National Open Data Agenda. The Open State Foundation, in association with other organisations, started Open Spending\textsuperscript{131} to implement a system in which all provincial authorities, local authorities and water management authorities use a common publication standard about their financial spending. Of all the open data published by the Dutch Government, the Ministry of Infrastructure and the Environment seems to have been most active in making data open.

Operating Environment

In the Netherlands, open data that can be accessed and re-used by everyone is data that:

From a Dutch perspective, open data is data that:\textsuperscript{132}

- Are paid for from the public purse and generated during or for the provision of a public service;
- Are available to the public;
- Are free of copyright and other third party rights;
- Are machine-readable and preferably comply with open standards (not PDF but XML, CSV etc.); and
- Can be used without restriction in the form of cost, compulsory registration, etc.

The Netherlands has established the National Data Portal\textsuperscript{133}, which provides an overview of the available data provided by Dutch government organisations. It is the main outlet on

\textsuperscript{131} Open Spending <https://www.openspending.nl>


\textsuperscript{133} National Data Portal <https://www.data.overheid.nl>
Dutch statistics. Ministries also publish data on several different websites, sometimes out of necessity. This has led to fragmented provision of open data.

While the Netherlands scores relatively well international benchmarks (it is currently number six on the Open Data Barometer Ranking), there is need for improvement in promoting open data. In October 2015, the Dutch Minister of the Interior, Ronald Plasterk in a letter\textsuperscript{134} to the Dutch Parliament made several commitments to open data. It is also evident that the Dutch Government is working towards promoting open data, holding the “view that relationships between the citizen and public sector authorities at all levels can and should be made (even) more open”.\textsuperscript{135}

Example of Open Data Projects, Outcomes and Applications

**Wikimuseums and Open Cultuur Data**

Open Culture Data is a network of cultural and heritage professionals, developers, designers, copyright specialists and open data experts, working on cultural heritage and encouraging the development of valuable cultural applications.

Wikimuseums had its initial workshop in Naples where members of Open Culture Data in conjunction with the city of Napoli showcased example of opening up of cultural and heritage data, as well as a marathon. The marathon used cultural and historical apps (Openstreetmap) that guided the user through important museums and areas of cultural importance. The objective was to bring the QRcode connected to Wikipedia closer to some of the places that had been looked at the previous day of the workshop. The Wikipedia entries for various cultural and historical points of interest in Napoli were amended and improved each day of the workshop.

Concluding Remarks

This paper analysed legislation, policy, regulatory settings, operational settings, roles and responsibilities for leading jurisdictions as identified in the Open Data Barometer. These

\begin{footnotesize}
\begin{enumerate}
\item\textsuperscript{134} <https://www.rijksoverheid.nl/documenten/kamerstukken/2015/11/30/kamerbrief-over-nationale-open-data-agenda-2016-noda>
\end{enumerate}
\end{footnotesize}
countries are: the United Kingdom, the United States, France, Canada, Denmark, the Netherlands, and New Zealand.

Observations suggest that lead jurisdictions see open government data not only about providing a transparent and accountable government, but also as a key economic driver in a data innovation society. Open data requires a cohesive, mandated responsibility to enable its execution which is ideally both centralised and localised as well as through citizen engagement and an appropriate legal framework.

As revealed in the communications with overseas entities, Open Data often has the greatest immediate impact for citizens at the sub-national levels. This is because applications and software developed as a result of Open Data at the sub-national levels often solve common problems which citizens can easily identify with. Common applications include those related to public transport, waste disposal, and zoning requirements. Future studies should seek to evaluate Open Data at sub-national levels, as well as evaluate how national Open Data frameworks are integrated with sub-national Open Data frameworks.

As seen across all jurisdictions, Open Data is still a developing concept with initial legislative, policy and regulatory work developed in leading jurisdictions, and implementation of policies and projects well under way. Leading jurisdictions and in particular the UK have experienced and addressed many of the barriers to Open Data operating within less developed nations. However, the UK’s legislative, regulatory and operational enablers have developed to address many of the initial barriers with new enablers including the recent passing of the *Digital Economy Act 2017*.

In practice the research has highlighted how diverse, inter-connected and context-specific each jurisdiction’s approach has been. In particular, it is clear that precisely because of the breadth of action some leading countries have taken it is difficult to isolate the particular contribution of any one element. However the existing legislative and policy settings have informed advances in open data in the jurisdictions examined, as have other operational tools and cultures of collaboration.

Considerable progress has been recorded and benefits delivered including those identified in the case studies highlighted throughout this article. There has, however, been very little work
done on measuring impact in any jurisdiction particularly from the dual limbs of social and economic savings together with the impact on participative democracy and citizen centric policy development and service delivery. Measuring impact of Open Data will be a critical component moving forward. Whether there is continued long-term investment in Open Data is dependent on its impact. Impact should be measured both in the short and long terms. Cost effectiveness, for example, may be slightly improved in the short term but over the long-term applications could have led to significant efficiencies for a department, and for an entire industry.

Additionally, there may not be the immediate desired impact from opening data as many innovators and researchers may not have had extensive previous experience for analysing data, and other researchers may not have experience using algorithms to run data analytics with machine-readable data. It may take time for some researchers and industry to gain the skills required to fully capitalise on open data. There has, however, never been a more exciting time for researchers with hundreds of thousands of previously privately held datasets opened up for research and observation. Likewise, the algorithmic tools required to run analytics are becoming more easy to use; one does not need a degree in statistics, data mining or computer science to perform these functions. With big data and open data, faster, more informed higher quality decisions can be made. Businesses have been empowered with tools and capabilities to better harness data to spot gaps in the market allowing for disruption of businesses. This same disruption is possible with government use of data, as well as researcher’s use of data.
Effectiveness in Mediation: a New Approach

Alysoun Boyle*

Abstract:

This Research Note reports the initial findings of content analysis of a selection of empirical studies of mediator techniques and their links with mediation outcomes. As part of the analysis, a new approach was devised for categorising effectiveness into simple and complex effectiveness, enabling a contextual comparative analysis of the selected literature. The analysis focused on contextual definitions of effectiveness in mediation. It was conducted as part of a larger research project seeking to establish what is known about effective mediators. The analysis findings suggest a lack of consistent definition and measurement of effectiveness in mediation research, the dominance of settlement as a measure of effectiveness in mediation across all contexts, and a lack of investigations of mediator influence over the achievement of mediation effectiveness. Finally, suggestions are made for future research, including investigation of possible links between complex effectiveness and the durability of mediated agreements.

Key Words: Mediation, effectiveness, definitions, content analysis

Introduction and Background

Effectiveness in mediation is often cited as being a core attribute, or selling point, of the mediation process. However, it has not always been clear which aspect of effectiveness is being lauded: its reputation for achieving settlement, or its reputation for achieving high levels of participant satisfaction, or both, or for some other beneficial attribute such as improving communication or supporting relationships.

This Research Note reports the results when a new approach, or analytical tool, is applied to measure the effectiveness in mediation. The analysis forms one component of a much larger research project that is seeking to establish what is known about mediator effectiveness. The larger project reports on an investigation of a selection of empirical

* PhD Candidate, The University of Newcastle. Author contact: Alysoun.Boyle@uon.edu.au.
studies of mediation (the selected literature), and includes bibliometric and terminological analyses, as well as a systematic review of the experimental methodologies applied to data collection within the selected literature.

This Research Note briefly describes the data to which the analysis is applied, the proposed new analytical tools, the findings of the analysis, and suggests some further research approaches to explore mediation effectiveness. In summary, the analysis confirms that a consistent methodology based on a theoretical framework approach is necessary to enable a productive contextual comparative analysis of mediation literature to be undertaken.

Four unexpected findings emerge from this work:

(i) There is no consistently applied definition of effectiveness in mediation, despite it being the most frequently applied unit of analysis in research about mediation;

(ii) Despite widespread claims from the mediation sector that participant satisfaction is a core benefit of mediation, the selected literature suggests that the majority of mediation research across a range of mediation contexts maintains a narrow focus on the achievement of settlement in mediation (the ‘simple’ effectiveness measure);

(iii) According to the selected literature, the two mediation contexts reporting the highest incidents of simple effectiveness are labour/management, and family; and

(iv) There is almost no investigation of the mediator’s influence on the achievement of simple or ‘complex’ effectiveness (which also includes a range of other impact variables) in mediation.

Findings from the analysis suggest that divergences in the definition and measurement of effectiveness in mediation might be linked to the subject mediation context. It is possible that there is a circular, self-reinforcing relationship between definition and context, where the context of the mediation influences the choice of definition and measurement of effectiveness, and that choice – and the subsequent research findings - reinforces established views about effectiveness within that context.

As a first step in the investigation of the possible relationship between context and definitions of effectiveness, this analysis resulted in the creation of a theoretical framework for exploring a contextual comparative analysis of simple and complex effectiveness in mediation.
Data for analysis

The selected empirical studies of mediation that have been the subject of the analysis were accessed through The American Bar Association (ABA) Section of Dispute Resolution Task Force on Research on Mediator Techniques (the Task Force), of which the author is a member. The Task Force compiled a selection of over 150 reports from the mediation literature in the US, UK, and Australia. To be included in the compilation, articles had to meet several criteria including reference to empirical investigations of the links between mediator techniques and mediation outcomes. The Task Force’s Final Report, including recommendations for future activity in the research field, has been finalised for publication by the ABA later in 2017.¹

Preferring to focus on primary sources, this research project applied further criteria to the Task Force compilation, in particular requiring that included articles be direct reports of empirical studies. Ultimately, forty-seven articles met the criteria for inclusion in this research project. Within this specific analysis, only thirty-eight of those articles referred specifically to effectiveness, and those thirty-eight (the selected literature) are the subject of this analysis.²

Simple and complex effectiveness in mediation

The concepts of simple and complex effectiveness were devised as part of this research project to enable comparative analysis of effectiveness between mediation contexts (contextual comparative analysis).

(i) Simple effectiveness

In this project, simple effectiveness refers to whether an agreement is reached within the mediation, and this appears to be the most frequently measured mediation outcome. For the purposes of this analysis, efficiency indicators and measures, such as timeliness and costs, are accepted as being qualifiers of settlement (e.g., where a settlement is achieved, was it achieved efficiently?), and are therefore included in simple effectiveness. Simple effectiveness is a case management statistic, a standard against which the effectiveness of many mediation services and programs is measured and assessed. Even a cursory assessment

---

² A list of the thirty-eight selected empirical studies is available from the author.
of the broader mediation literature suggests that the focus on simple effectiveness may have influenced the design of many mediation programs and services.\(^3\) So widespread is its application in research and practice, simple effectiveness could be said to be an “industry standard” for mediation across all dispute and mediation contexts.

A focus on simple effectiveness bypasses more complex analyses of what actually happens within mediation, such as the contributions and influences of the mediator.

(ii) Complex effectiveness

Complex effectiveness includes the measurement of any, or a combination of any, mediation outcomes additional to the achievement of settlement,\(^4\) including, but not limited to, the following:

- Participant satisfaction (accepted as incorporating factors that relate to perceptions of fairness),
- Rates of compliance,
- Nature of agreements, and
- Improvement in participants’ post-dispute relationship.

It could be said that the above measures are each relevant to overall participant satisfaction and perceptions of fairness. In this Note, they are referred to collectively as “fairness and satisfaction measures”. Some of the selected studies used measures that are additional to, rather than being components of, effectiveness; for example, one program evaluation report includes measures of effectiveness drawn from the ADR objectives defined by the Australian Law Reform Commission in 1998.\(^5\) Where relevant, the additional measures are specified in this analysis.

Because of the nature of its achievements, complex effectiveness could be said to achieve broader resolution rather than narrow settlement.

---


\(^4\) None of the selected literature investigated effectiveness that did not include the achievement of settlement.

Complex effectiveness has been said to produce more durable agreements than simple effectiveness, yet is less likely to be considered than simple effectiveness in the assessment of mediation services and programs, and appears less frequently in the mediation literature as a component, or measure of effectiveness. In addition, while mediators may play a role in achieving the outcome standard of simple effectiveness, their role in relation to complex effectiveness is largely unexplored.

Method

The thirty-eight selected empirical studies of mediation were subjected to an initial manual content analysis to determine the context of each study, and to clarify the definition and measures of effectiveness that were applied in each individual study. Because of the inconsistencies in the definitions and measures relevant to the effectiveness component, the analysis was somewhat cumbersome. Subsequently, the measure of simple or complex effectiveness was devised and applied to a second analysis of the selected literature; the results of that analysis are reported below. The selected mediation literature includes articles that report on recent studies as well as studies conducted more than 30 years ago. In some instances, historical contexts may differ slightly from their modern counterparts. For example, in the US in the 1970s, community mediations were often conducted through a small number of Neighbourhood Justice Centers and the types of disputes to be resolved were limited to minor disputes between neighbours, landlord/tenant disputes, minor consumer complaints, and family disputes.

A modern understanding of “community mediation” in the United States (U.S.) may have changed to include a broader category of disputes (such as consumer complaints), but not sufficiently to affect the focus of this analysis.

Mediation Context

The initial analysis of the selected literature established seven contexts within which subject mediations were conducted or within which research or evaluation took place; each of the selected thirty-eight reports was allocated to one context. Within this analysis, context is

---


taken to mean the setting within which the investigated mediations occur, as reported in each study, or as deduced during the analysis. The context groupings are described below:

- Court-connected context: as described in the selected empirical studies, a court-connected mediation is taken to be one that is conducted within, or in association with, a court or tribunal, and/or conducted by a practising or retired judge or legal practitioner. This contextual category does not include assessments or evaluations of court-connected mediation programs or services, which are included in a specific category of their own;

- Evaluation context: Funded evaluations or assessments of existing mediation programs and services where program and service evaluations tend to have their evaluation criteria established as part of their terms of reference, and such criteria usually include readily quantifiable measures of efficiency such as settlement rates, timeliness, and cost reductions;

- Labour-management context: as described in the selected empirical studies, a labour-management mediation is taken to be one where the issues in dispute concern collective workplace conditions, and the parties are management and employees, or their representatives and does not include workplace mediation between individual employees and individual employers/managers;

- Community context: as described in the selected empirical studies, a community mediation is taken to be one that is conducted at a community mediation centre, regardless of the professional standing of the mediator or the nature of the dispute;

- Family/divorce/child custody context: as described in the selected empirical studies, a family/divorce/child custody mediation is taken to be one in which the mediation is conducted by a family mediation service, and/or concerns matters associated with divorce proceedings and/or child custody and visitation rights;

---

8 Family/divorce/child custody mediations usually occur in association with a purpose-specific court; however, they differ from legal mediations in that, in the selected literature, they are not reported to routinely include judge/legal practitioner mediators.
• Simulated mediation context: in this small cohort of studies, evaluation criteria were reported as having being set as part of the simulation design in the laboratory setting; and

• Construction and building context: there is only one report from a study of mediations conducted in a construction setting, and the majority of the mediators were reported to be either lawyers or retired judges; despite the legal background of the mediators, this study has been differentiated from legal and court connected mediations because of the nature and setting of the disputes themselves.

After each study had been allocated to its relevant context category, the definitions and measures of effectiveness in all studies were analysed according to the measure of simple and complex effectiveness.

**Results of contextual comparative analysis**

The results of the analysis suggest that applying simple and complex effectiveness criteria can be productive in contextual comparative analyses of effectiveness in mediation. The analysis shows a clear dominance of simple effectiveness across most contexts, especially in labour/management, family, and court-connected mediations. Although the use of complex effectiveness measures was reported in a significant minority of the studies, it is surprising that complex effectiveness is not applied more widely in empirical studies of mediation.

In summary, of the thirty-eight selected journal articles that explore issues related specifically to effectiveness, twenty-three, or around 60%, refer to settlement (including efficiency measures) as the sole measure for effectiveness. The two contexts with the highest proportion of simple effectiveness measures are the labour-management context (100%), and the family/divorce/custody context (80%). Of the legal context studies, 70% measured simple effectiveness only.
The results of the contextual comparative analysis are listed below, in descending order of contextual prevalence in the selected mediation literature.

(i) The court-connected context

Of the ten studies of mediations reported to have been conducted in a court-connected context,9 seven reported only on simple effectiveness.10 The other three studies reported on complex effectiveness that was measured in terms of settlement plus fairness and satisfaction measures.

(ii) Mediation program and services evaluations and assessments


10 Where this analysis reports that effectiveness was measured in terms of whether or not settlement was reached, this binary question is considered to be different from so-called “settlement rates” because the studies do not consistently compare numbers of settlements across studies, or with other benchmarks.
Eight reports in the selected literature are of funded assessments or evaluations of mediation programs or services.\textsuperscript{11} This category necessarily includes several mediation contexts, but program evaluations and assessments differ in that their research and evaluation criteria are established in their terms of reference. The terms of reference are likely to include readily quantifiable factors indicating effective case management: settlement, timeliness, and cost reductions.

The eight evaluations and assessments of mediation programs and services include five evaluations of court-connected mediation programs, one of a workplace mediation program, one of mediation services in the financial sector, and one of mediation services in a family/divorce/custody context.

Of the eight reports in this category, two report on simple effectiveness. The other six report on measures of complex effectiveness including measures of fairness and satisfaction. One of the six includes an additional measure of participant perceptions of justice within the subject (workplace) mediation program. Another includes additional measures of effectiveness that drew from ADR objectives defined by the Australian Law Reform Commission in 1998.\textsuperscript{12}

The three evaluations of court-connected programs/services that measure complex effectiveness were conducted over a period of eight years, by the same researcher; that researcher also co-conducted a fourth program evaluation in a different context, the latter including measures of complex effectiveness. Because it is not clear whether the funding bodies, the researcher, or some other factor influenced the methodology and terms of reference for the five evaluations, it cannot be ascertained if they represent a trend (especially in Australia) towards measures of complex effectiveness in evaluations of court-connected, and other, mediation programs.


\textsuperscript{12} Australian Law Reform Commission, above n 5.
(iii) **The labour-management context**

Of the seven studies of mediations reported to have been conducted in a labour/management context, all seven report only on simple effectiveness. None define or measure any aspect of complex effectiveness. One of the reported studies does include measures of the percentage of issues resolved, and any observed movement in the parties’ positions or concessions made, but these relate directly to the ultimate achievement of settlement.

(iv) **The community context**

Of the five studies of mediations reported to have been conducted in a community context, two report on simple effectiveness, and three on complex effectiveness using fairness and satisfaction measures.

(v) **The family/divorce/custody context**

Of the five studies of mediations reported to have been conducted in a family, divorce, or custody context, four report on simple effectiveness. One study defines effectiveness in

---


terms of settlement plus fairness and satisfaction measures, including the durability of the agreement and the participants’ attitudes to the mediator.

(vi) Simulated mediations

Two reports are of studies conducted using simulated dispute scenarios and mediation sessions.\(^{16}\) The two reports measure effectiveness only in terms of settlement. Additional factors, such as participant perceptions of the mediator, participant use of reframing, and participant assessment of the mediator are the more dominant focus of the studies.

(vii) Construction and business context

Only one of the studies investigated mediations conducted in the construction and business context.\(^{17}\) It defines effectiveness in terms of whether or not settlement was reached, plus additional efficiency measures (reduction in costs, and timeliness).

Commentary

The application of simple and complex effectiveness measures has been productive in this analysis of empirical studies of mediation conducted across a range of contexts. Using simple and complex effectiveness as an analysis tool has enabled an overall analysis of

---


approaches to effectiveness, as well as contextual comparative analysis where previously such analysis had been too cumbersome to undertake productively.

Fifteen of the thirty-eight selected studies do include measures of complex effectiveness (i.e., around 39%). Additional year-of-publication analysis of the selected literature suggests that in no decade since the 1970s have measures of complex effectiveness outnumbered measures of simple effectiveness.\(^{18}\) The same year-of-publication analysis of the selected literature suggests that the 1990s might have been the decade in which there was greatest application of complex effectiveness measures.\(^{19}\) However, caution is advisable with this year-of-publication analysis because the very small number of studies published in some decades in itself precludes valid interpretation. Year-of-publication data alone cannot take into account the range of factors that might influence the choice of effectiveness measures at any time, in any mediation context.

An unintended finding from this analysis is the apparent lack of consistently applied definitions and measures of effectiveness in mediation. The application of the analysis tool (simple and complex effectiveness) has confirmed the dominance of simple effectiveness as a definition and unit of analysis in mediation research across many mediation contexts. Conversely, the analysis has revealed the minority status of complex effectiveness in mediation research, despite its relevance to participant satisfaction and its potential links to agreement durability. It is not known to what extent the focus on simple effectiveness limits the scope for assessment and development of mediation programs, the scope of research into mediation, and developments in the practice of mediation.

Although limited to the selected literature, the results of this analysis reveal which mediation contexts appear to be most strongly focused on achieving settlement, and which tend to take into account participant perceptions. Useful further research could include applying simple and complex effectiveness to analysis beyond the selected literature.

A small number of studies in the selected literature do report on investigations of the durability of mediated agreements, but none investigated the potential links between simple or complex effectiveness and agreement durability. A small number of publications have

\(^{18}\) For example, only one of the selected articles was published in the 1970s, and only two were published in the 2010s.

\(^{19}\) Of the selected literature, none of the seven articles published in the 1980s included any measures of complex effectiveness; in the 1990s, a little over 69% included complex effectiveness; in the 2000s, 60% included measures of complex effectiveness.
considered possible links between complex effectiveness and agreement durability; however, they are non-empirical studies that are not included in the selected literature.

Conclusion

As part of a broader research project, this analysis proposes the application of simple and complex types of effectiveness to overcome the divergent definitions and measurements of effectiveness in mediation, and to facilitate contextual comparative analysis of the selected empirical studies of mediation. Simple effectiveness includes only the measurement of whether settlement is achieved in the mediation; complex effectiveness includes several additional factors, usually relating to perceptions of fairness and satisfaction, in addition to whether settlement is achieved. Analysis of the selected literature demonstrates the dominance of simple effectiveness measures across seven mediation contexts. Unexpectedly, it suggests both a lack of consistently applied definitions and measures of effectiveness in mediation, and the dominance of settlement as a unit of analysis effectiveness.

This analysis has revealed a surprising lack of investigations into the mediator’s influence on the achievement of either simple or complex effectiveness.

Accepting that durable mediation agreements are valuable for a range of reasons, additional empirical studies should investigate the reported links between the durability of mediation agreements and the achievement of complex effectiveness, as well as the influence of the mediator. The outcomes of such research would be major contributors to discussion about how mediation offers a return on investment.

Limitations

The relatively small number of selected empirical studies is a limitation; however, this reflects the relatively small and specialised field that is explored in mediation research; it also reflects the criteria for inclusion in this research project, in particular that a report be focussed on an empirical investigation of the links between mediator techniques and mediation outcomes. A second potential limitation is that the majority of included articles report on empirical studies conducted in the US. Although this is a reflection of the membership and affiliation of the Task Force, it is also a reflection of the US’ dominance in the field of mediation research; however, a smaller number of the articles do report on studies conducted in the UK, Australia, and the Netherlands. The selected studies do not include empirical studies from Asia, or broader Europe, or other regions; nor do they include the so-called
“grey literature” which, though a potential source of valuable research findings, is unavailable through public forums.

A recent article published in 2017 reports on an empirical investigation of what mediators do to facilitate successful, or effective, mediations. That investigation collected data from 13 mediators practising in France, Spain, Luxembourg, and Canada, but considers only simple effectiveness. Despite being published almost 40 years after the oldest report in the selected literature, the 2017 report confirms the literature’s dominant preference for defining effectiveness in mediation only by the achievement of agreement.

---