addresses
How Rights Change: Freedom of Speech in the Digital Era
Jack M Balkin
5

Banes of an Income Tax: Legal Fictions, Elections, Hypothetical Determinations, Related Party Debt
H David Rosenbloom
17

articles
Refugee Law: The Shifting Balance
Ronald Sackville
37

Judging Refugees: The Clash of Power and Institutions in the Development of Australian Refugee Law
Mary Crock
51

First Head Revisited: A Single Industrial Relations System under the Trade and Commerce Power
David McCann
75

Being Ms B: B, Autonomy and the Nature of Legal Regulation
Derek Morgan & Kenneth Veitch
107

case commentary
A Prelude to the Demise of Teoh: The High Court Decision in Re Minister for Immigration and Multicultural Affairs; Ex parte Lam
Wendy Lacey
131
Faculty of Law

Dean
Ronald McCallum, Bjur LLB Monash, LLM Qu, Blake Dawson Waldron Professor of Industrial Law

Pro Dean (Staff Development)
Barbara McDonald, LLM Lond, BA LLB, Associate Professor (Semester 1)
Julie Stubbs, BA W'gong MA Tor, Associate Professor (Semester 2)

Pro Dean (Teaching Programs)
Patricia Loughlan, BA LLM Tor, PhD, Associate Professor (Semester 1)
Donald Rothwell, BA LLB Qld, LLM Alta, MA Calg, PhD, Associate Professor (Semester 2)

Associate Deans
(Undergraduate)
Graeme Coss, LLB, LLM (Hons) Syd, Grad Dip Inf&LibStud Curtin, Senior Lecturer

(Postgraduate)
Jennifer G Hill, BCL Oxf, BA LLB, Professor (Research)
Roger Magnusson, BA, LLM ANU, PhD Melb, Associate Professor (Coursework)

International Students
Bernhard W Boer, BA LLM Melb, Professor of Environmental Law (Personal Chair)

Lee Aitken, BA, LLM ANU, BCL Oxf, Associate Professor
Margaret Allars, DPhil Oxf, BA LLB, Professor
Ross Anderson, LLM Lond, LLB, Senior Lecturer
Patricia Apps, BArch UNSW, MED Yale, PhD Camb, ARAIA, Professor in Public Economics in Law (Personal Chair)
Emma Armson, BEc, LLB Macq, LLM UNSW, Lecturer
Hilary Astor, BTech(Law) PhD Brunel, Abbott Tout Professor of Litigation and Dispute Resolution
Belinda Bennett, BEc LLB Macq, LLM SJD Wisc, Associate Professor
Celeste Black, BA Harv, JD Penn, LLB, Lecturer
Fiona Burns, BA LLB LLM Syd, LLM Camb, PhD ANU, Senior Lecturer
Lee Burns, BComm LLB UNSW, LLM, Sesqui Associate Professor in Taxation Law
Peter Butt, BA LLB Syd, Professor
Terry R Carney, LLB DipCrim Melb, PhD, Monash, Professor
John W Carter, PhD Camb, BA LLB, Professor in Commercial Law (Personal Chair) (fractional)
Judith Cashmore, BA, DipEd, Adel, MED Newcastle, PhD Macquarie, Research, Level B Academic (fractional)
Graeme Cooper, BA, LLB, LLM Syd, LLM Illinois, LLM, JSD Col, Professor of Taxation Law
Mary Crock, BA LLB PhD Melb, Senior Lecturer
Christopher Cunneen, BA DipEd UNSW, MA, Associate Professor
Kristen Daglish, BA, Macq, LLB(Hons) Macq, MIL Syd, Associate Lecturer (fractional)
Bernard Dunne, BA LLB Macq, Lecturer
Mark J Findlay, BA LLB ANU, DipCrim MSc Edin, LLM, Professor
Nicola E Franklin, BA Natal, LLB Natal & Camb, DipComparLegalStud Camb, Senior Lecturer
Saul Fridman, LLB W Ontario, BCL Oxf, Senior Lecturer
Jennifer Gage, BEc, LLB Syd, Lecturer (fractional)
Peter Gerangelos, BA, LLB Syd, LLM UNSW, Lecturer
Reg Graycar, LLB Adel, LLM Harv, Professor of Law
Helen Irving, BA Melb, Etudes Libres (License) Lausanne, MPhil Camb, PhD, Associate Professor
Fleur Johns, BA, LLB Melb, LLM Harv, Lecturer
Isabel Karpin, LLM Harv, SJD Col, BA LLB, Senior Lecturer
Abigail Kohn, BA Harv, MPhil (Criminology) Camb, PhD (Medical Anthropology) Calif, U2000 Postdoctoral Fellow, Level A Academic
Ghena Krayem, BA, LLB Syd, Associate Lecturer (fractional)
Rosemary Lyster, BA LLB LLM Natal, Senior Lecturer
Gail Mason, LLB, Qld, DipCrim Melb, MA SUNJ, PhD LaTrobe, Senior Lecturer
Les McRlimmon, BA LLB Alta, LLM Qld, Associate Professor
Rebecca Millar, BSc Auck, LLB UNSW, Senior Lecturer in Taxation
Jenni Millbank, LLM Br Com, BA LLB, Senior Lecturer
Les Nottage, BCA, LLB VUW, LLM Kyoto, Senior Lecturer
Patrick N Parkinson, MA Oxfl, LLM Ill, Professor
Elisabeth Peden, BA, LLB Syd, PhD Camb, Senior Lecturer
Kristin Savell, BSc, LLB Macq, LLM McGill, PhD Camb, Lecturer
Belinda Smith, BEc, LLB Syd, Lecturer
Tim Stephens, BA(Hons) LLB(Hons) Syd, MA UTZ, Associate Lecturer (fractional)
Gerry Bates, LLB, PhD Birm, Senior Lecturer (Environmental law)
Christopher Birch, BA LLB PhD (Aspects of Legal Reasoning) (S1 & S2)
Anne McNaughton, BA N’cun, LLB UNSW, LLM Tübingen, LLM UTS (Legal System of the European Law (S1))
Peter B Shea, BHA GradDip(HealthAdmin) UNSW, DipEnvStud Macq, MB BS MPH DPM DipCrim Syd, FRANZCP, FRACMA LHA AFAIM (Challis Lecturer in Forensic Psychiatry) (Forensic Psychiatry)
Brian Preston, BA LLB Macq (Biodiversity Law)
Robin H Woellner, BA LLB (Taxation Administration)
Address
How Rights Change: Freedom of Speech in the Digital Era†
JACK M BALKIN*

Abstract

Technological change produces new forms of social conflict. The digital revolution is no exception: lowering the costs of distribution and content creation inevitably creates conflicts between ordinary individuals and the information industries. Freedom of speech is a key site for these struggles, as media companies repeatedly attempt to expand intellectual property rights at the expense of individual free expression, while simultaneously invoking freedom of expression to oppose telecommunications regulation. This stunted vision of free speech undermines the creative and participatory possibilities of the digital age; it treats ordinary individuals as passive consumers rather than active producers of their cultural world. We must promote a democratic culture that celebrates interactivity and widespread cultural participation. Earlier free speech theories concerned with democratic deliberation were adapted to the twentieth century world of broadcast media, in which only a relatively few people controlled access to mass communication. Free speech theory must now be dedicated to promoting each individual's ability to participate in the growth and development of culture.

* Knight Professor of Constitutional Law and the First Amendment, and Director, The Information Society Project, Yale Law School.
I am honored to participate in a lecture series dedicated to a distinguished figure in 20th Century jurisprudence, Julius Stone. Much of Stone’s work was concerned with law in its social context, and my remarks are very much in that spirit. My topic is how freedom of speech changes in the digital age, and, more generally, how disputes about rights respond to changes in economic, social, and technological circumstances.

Technological change, I shall argue, spurs social conflict, which is fought out in law and in the discourse of rights. The new digital technologies change the social conditions of speech. They create new forms of contention between ordinary individuals, who now possess tremendous new opportunities to communicate and create, and the information industries, who want to expand markets and maximise profits from the same technologies. These conflicts will be fought out in debates over the free speech principle. Hence, in light of these changed circumstances, we must pay careful attention to the goals of freedom of speech in the digital age. In my view, the point of free expression is to promote a democratic culture, a culture in which ordinary individuals are free to create, innovate, and participate in the processes of meaning-making that in turn constitute them as individuals.

In studying the Internet, or indeed any profound technological change, we shall immediately get off on the wrong foot if we ask what is genuinely new, for almost every change has some precedent in human history. Rather, the key question is not novelty but salience: what elements of the social world does a new technology make particularly salient that went relatively unnoticed before? What features of human activity or of the human condition does technological change foreground, emphasise or problematise? And what are the consequences for human freedom of making this aspect more important, more pervasive, or more central than it was before? These features of social life may well have always been present, but now they appear to us with a special sense of importance or urgency. This phenomenon is true, I think, of all technological change. Technological change modifies and disrupts social relations. It foregrounds certain elements and aspects of social life, making them more central, more salient, more important than they were before.

The digital revolution has reduced the costs of copying and distributing information drastically, almost to the vanishing point. Indeed, one sends digital information over the Internet by making copies of it. Lowering the costs of copying and distributing material has many important effects: it makes it easier for people to talk to each other; it makes it easier to send things across geographical borders; it makes it possible for new communities of interest to form, and older communities to gain new members.

A second important feature of the digital age is the development of common standards of information exchange. The Internet provides common standards and protocols for moving information from one place to another. Microsoft’s operating systems have provided common standards for creation of software. The industrial
age greatly benefited from the creation of interchangeable standardised parts; in like fashion, common standards for codifying and transmitting information have been crucial to the spectacular successes of the digital age. Common standards are important for another reason: They also make it easier to alter information; they facilitate the development of software programs that allow people to modify information coded according to a standard format. So the digital age not only lowers the costs of distribution, but also the costs of content creation, annotation, and modification.

How do these features of the digital age change the social conditions of freedom of speech? Lowering the costs of transmitting, distributing, creating, and modifying information has important democratising and decentralising effects. For it also lowers the costs of forming communities of interest, of interacting with people, of making new things out of old things, of innovating, copying, altering and modifying, of distributing an individual’s ideas to large numbers of people.

The mass media that developed in the 20th Century are asymmetrical — one entity speaking to many persons, and they are unidirectional — the broadcaster sends information to you through the radio or television, but you cannot use the radio or television to talk back. These mass media also serve as bottlenecks and gatekeepers. Suppose you have a good idea. Can you simply decide to put it on TV? No, you must go through a whole set of intermediaries — broadcast companies, producers and so on. The Internet is quite different. It is neither asymmetrical nor unidirectional; lots of people can broadcast and talk back to each other. Equally important, the Internet allows you to route around the intermediaries and gatekeepers of the traditional mass media. You can publish your book on the Web. You can make your own movie or demo tape and distribute it on your website. You can say whatever you want on your own weblog.

Many people assumed that the Internet would displace the mass media and publishing houses as traditional gatekeepers of content and quality. This has not occurred. Traditional mass media have remained quite robust. Rather, the Internet has provided an additional layer of communication that rests atop the mass media, draws from it, and in turn influences it.

Internet speech has two important characteristics: It routes around traditional mass media and it gloms onto it. To route around means to avoid the gatekeepers and bottlenecks of the mass media, to do an end run around them. To glom on is to appropriate or take something from mass media and make use of it. When people use the word ‘appropriate’, they often mean taking something for one’s exclusive use. But by ‘glomming on’ I mean non-exclusive appropriation of content. Information goods are particularly suitable for non-exclusive appropriation because they are non-rivalrous: I can have an idea and you can have it too. I can incorporate elements of something I have seen or heard in my work and you can too. When we say that people glom on to the products of the mass media, we mean that people use what they find in the mass media as a platform for annotation and commentary, and that they use what they find as raw materials for innovation and creativity. Routing around and glomming on are the two most important features
of the relationship between the Internet and the mass media: People route around the traditional gatekeepers of the mass media, but they also glom onto aspects of mass media, and use them for their own purposes. Indeed, routing around and glomming on are the two most important features of Internet speech generally.

Let me offer a few examples. The first concerns weblogs, or blogs as they are usually called. Blogs represent a further reduction in the cost of publishing, allowing people to instantaneously publish their thoughts and opinions in journal form. Entire communities of people, called bloggers, now use these weblogs, and they write about everything under the sun, from politics to popular culture, to detailed accounts of their love lives, to what their cat did the other day. They link to each other, respond to each other, and criticise each other. Many bloggers comment on contemporary politics and public issues and on how the mass media are covering them. These bloggers link to and discuss stories that appear in the New York Times, the Washington Post, and other mainstream journals that run websites. Still other bloggers link to and comment on sources that aren’t as well known and they bring these sources to the attention of a wider audience.

Bloggers route around the traditional mass media because people can go directly to the blogger’s website. But bloggers also glom on: they appropriate quotations and information from the mass media, and use them to offer their commentary and make their points. Other bloggers quote, build on, and criticise what these bloggers have to say, and in this way ideas spread throughout what is called the blogosphere. Blogs have become a supplementary form of journalism that some mainstream journalists turn to for story ideas. Some bloggers are journalists who post stories that they cannot get published in the mainstream press. Others are simply amateurs who have opinions. But their journalism has a feedback effect on the mainstream media. Sometimes bloggers will keep a story alive for days or even weeks until mainstream journalists pick it up and run with it. Thus, through routing around and glomming onto the mass media, blogs interact with mass media and affect and influence it.

My second example is a website called Television Without Pity,1 run by a group of (mostly) Canadian television fans. They watch television — mostly American television — and offer play-by-play accounts of what happened on each episode, including large swaths of the actual dialogue, along with their own witty and sarcastic commentary. Here again we have both routing around and glomming on: the people at Television Without Pity reach their audiences directly without going through the mass media as professional television critics do. Indeed, because of the nature of their criticism it’s unlikely that they could ever be hired as professional television critics; but they don’t depend on traditional media gatekeepers. Television Without Pity involves lots of glomming on, because their critics make considerable use of plots, characters and dialogue to make their points.

---

1 <http://www.televisionwithoutpity.com>
Finally, as with blogs, Television Without Pity has had interesting feedback effects on the mass media it routes around and gloms onto. Television producers want to know what viewers think about their shows — that’s why they have focus groups and do survey research. Several television producers became quite interested in Television Without Pity: Here was a group of highly articulate viewers who were telling the world what they thought about these shows, and who had their own audience of fellow viewers that they in turn influenced. The Internet created a way for viewers to talk back to the people who produce television shows, a method that was far more immediate and interactive than traditional forms like fan mail.

My third example is fan fiction: People write stories involving characters in their favorite books, movies, or television shows, carrying existing plot lines further and sometimes constructing entirely new episodes. There is a tremendous amount of fan fiction on the Internet. In particular, there seem to be an enormous number of people who like to write stories about the characters from the television series, *Star Trek*. The number of homosexual romances between Captain Kirk and his Vulcan first officer, Mr Spock, is probably beyond counting at this point. Of course, fan fiction is not limited to television shows and movies. There are any number of fan fiction stories about Harry Potter, in which Harry Potter has all sorts of adventures, and does things that would probably make JK Rowling’s hair stand on end. Once again we see the characteristic features of Internet speech — ordinary people routing around traditional mass media gatekeepers, glomming onto the products of mass media, appropriating elements of what they find, using them as launching pads for innovation and imagination, and turning them to their own creative purposes.

My fourth and final example involves one of the most famous instances of glomming on: The Phantom Edit. You may have heard of the *Star Wars* series of movies, and you may know that Episode I, which is actually the fourth picture in the series, is called *The Phantom Menace*. A *Star Wars* fan got a copy of *The Phantom Menace* in digital format, and reedited it digitally. (As I mentioned previously, common standards for digital distribution also mean that there can be common standards for digital manipulation.) Now this fellow took a particular dislike to one of the movie’s characters, Jar Jar Binks, a strange, buffoonish sort of alien with very long ears and nose, who talks in a sort of faux-Caribbean patois. The fan edited all the parts of the movie that featured Jar Jar Binks, digitally altering the screen so that Jar Jar is missing as much as possible. The Phantom Edit showed that if you had a digital copy of the movie, you could essentially remake the movie; you could do your own director’s cut. When George Lucas, the director of *The Phantom Menace*, found out about The Phantom Edit, he was not pleased, for obvious reasons. But The Phantom Edit is another way of talking back, if you will, to a form of mass media that was, from its very earliest days, asymmetrical and unidirectional. It is not the passive consumption of a media product by a consumer. Rather, it involves a viewer actively producing something new through
digital technologies. It exemplifies what the new digital technologies make possible: ordinary people using these technologies to comment on, annotate, and appropriate mass media products for their own purposes and uses.

Digital technologies foster interactivity; they allow ordinary people to route around traditional media gatekeepers and offer new ways of appropriating and transforming what people find in the mass media. But here is the catch. These same features of the new technologies that empower ordinary individuals also create a very powerful and serious social conflict. That should not be surprising. We often think of new technology as something that liberates us, if we are optimists, or threatens us, if we are pessimists. Technology produces either utopia or dystopia. But what technology more often does is create social conflict. It empowers us with respect to others and makes us vulnerable to others in new ways. Technology mediates and reconstructs our relationships to other people; it produces and redistributes power and vulnerability.

Technological change never simply empowers or simply endangers people. It changes their relations to others: their social relations, their economic relations, their cultural relations, their relations of power and authority, and so on. Thus, technological change creates new forms of social conflict because it allows human power to be exercised and distributed in new ways. It creates new communities of interest and divides old ones; it pits existing groups (and newly formed communities of interest) against each other in new ways.

The Internet and digital technologies open up a conflict between the expanding information industries, who make money from digital technologies, and the ordinary people, or ‘end users,’ who surf the Internet. The digital revolution is both a social revolution and an economic revolution. Lowering the costs of information production and distribution opens up new markets, allows businesses to reach more people and creates new opportunities for making money. Information and information products become an increasingly important form of wealth in society, and an increasing source of economic power and influence. Unsurprisingly, information industries seek to maximise their control over distribution networks for digital content, and to maximise the value of their investments in intellectual property.

But the very technology that allows businesses to distribute their technology and make money out of it also allows other people to copy it and change it. The same technologies that lower costs of distribution and content production also make it easier for end users to copy, distribute, manipulate, and appropriate information. The same digital technologies that allow George Lucas to make The Phantom Menace allow a Star Wars fan to make ‘The Phantom Edit,’ and to copy and distribute it to other people easily and costlessly.

This is the central conflict of the digital age: On the one hand, we have a set of technologies that open up new possibilities for communication, creativity, and innovation, decentralising control over information and democratising access to audiences. On the other hand, we witness the increasing importance of information
as a commodity to be bought and sold, and the expansion of new markets for the sale and development of intellectual property and media products. These two effects, caused by the same technological advances, rapidly come into conflict with each other. The very same technologies that create new possibilities for general democratic cultural participation clash with the desire to exploit new markets and accumulate wealth.

Like many social conflicts, this one is fought out in law, and, in particular, through the meaning of the free speech principle. The conflict arises in several different locations in legal doctrine. Here I will mention only two. The first is intellectual property; the second is telecommunications policy.

As intellectual property became an increasingly important source of wealth in the 20th Century, businesses pushed courts and legislatures to expand the boundaries of intellectual property law in two directions, horizontally and vertically. Media corporations have sought to expand intellectual property rights horizontally by including protection for derivative works, sequels, characters, plots, and so on, and vertically, by increasing the length of intellectual property protections like copyright terms.

Media companies, however, have not limited themselves to legal devices. They have also turned to technology to protect their intellectual property interests. A central example is digital rights management schemes, technological devices that prevent copying of and control access to digital content. In the United States, the Digital Millennium Copyright Act 1998\(^2\) created a new species of legal rights, sometimes called 'paracopyright,' that make it unlawful to circumvent these technological devices or distribute circumvention devices to others. Although digital rights management is often justified as a means of preventing unauthorised copying, it actually goes much further. It is part of a general strategy of control over access to digital content, including digital content that has been purchased by the end user. Digital rights management schemes, for example, can make digital content unreadable after a certain number of uses; they can control the geographical places where content can be viewed, they can require that content be viewed in a particular order, they can keep viewers from skipping through commercials and so on. Paracopyright creates legal rights against consumers and others who wish to modify or route around these forms of technological control.

Matters have come to a head as copying and modification of digital content have become widespread, and media companies have sought in increasingly aggressive ways to protect their existing rights and expand them further. The problem is that some of these legal and technological strategies are seriously curtailing freedom of expression. Not surprisingly, media companies have generally resisted the idea that freedom of speech limits the continued expansion of their intellectual property rights. Nevertheless, at the same time media

corporations have avidly pushed for constitutional limits on telecommunications regulation on the ground that these regulations violate their free speech rights. We might call this combination the ‘capitalist theory’ of freedom of speech.

From a free speech standpoint, there is something odd about this agenda. It seems to expand free speech rights in telecommunications law, while contracting them in intellectual property. But in fact there is a deeper unity here. The capitalist theory does not really promote the values of freedom of expression so much as it employs concepts of speech and property selectively to promote the economic interests of the most powerful economic entities that characterise the digital age — the media corporations that produce and sell media products and other informational goods. Promoting freedom of speech as an anti-regulatory device in telecommunications law while denying it any force as a corrective to the growth of intellectual property rights is a theory of freedom of expression tailor-made for the information industries and the protection of their investments in capital.

This is hardly the only possible way to think about what freedom of speech might mean in the digital age. Indeed, this approach largely undermines the participatory promise of the digital revolution. To a considerable extent, it seeks to reassert an earlier conception of the end user as essentially a passive recipient of the media goods produced by media companies. The point of controlling distribution networks and rigidly controlling intellectual property rights is to promote consumption rather than production and individual creativity, to arrange matters so that the end-user is placed in a sort of consumerist utopia, continuously being offered a series of opportunities to consume or buy, which are seamlessly melded with other forms of communication. In this vision’s most perfect form, communication and consumption would become one.

There is nothing wrong in theory with maximising buying opportunities, but it is important to understand that this is not the same thing as a system of freedom of expression. Rather, it is a system of freedom of consumption, where liberty means freedom to choose among media goods. That is a stunted vision of free expression because it undermines the creative and participatory possibilities of digital technologies.

I’d like to offer a different notion of freedom of speech for the digital age: I believe that the point of freedom of speech is to promote a democratic culture. What is a democratic culture? It is a culture in which people can participate actively in the creation of cultural meanings that in turn constitute them. A democratic culture is democratic not in the sense that everyone gets to vote on what is in culture. It is democratic in the sense that everyone gets to participate in the production of culture. People are free to express their individuality through creativity and through participation in the forms of meaning-making that, in turn, constitute them and other people in society.

The last great change in communications technology was the emergence of the broadcast media. But only a relatively small number of people had the power to speak using these technologies. This social organisation of speech created a
genuine problem for the system of free expression, because not everyone was able
to use the technology equally. People were able to listen, but not to be broadcasters
themselves. That problem led to free speech theories that made a virtue out of
necessity: They argued that the goal of freedom of speech is providing information
necessary for democratic self-governance, and particularly democratic
deliberation about public issues. The great philosopher of education, Alexander
Meiklejohn, was probably the most famous advocate of this view. Freedom of
speech did not exist to promote individual autonomy, he argued. Rather, the
purpose of free speech was to promote democratic deliberation about the great
issues of the day. As Meiklejohn famously put it, it is not important that everyone
gets to speak but that everything worth saying gets said.3

Democracy-based theories responded to the social conditions of speech
produced by the broadcasting technologies created in the 20th Century: a world in
which a relatively small number of people controlled radio and television
broadcasting, and later cable and satellite broadcasting. Free speech theorists
worried that democratic discourse would be skewed and that the information
necessary for wise governance would be impaired when the most powerful
broadcasting entities were held in a relatively small number of hands. This concern
justified public interest regulation of broadcasting, cable, and other mass media.

Because mass media will remain a central feature of culture for the foreseeable
future, some structural public interest regulation continues to make sense,
particularly as mass media consolidate and exercise their economic power
globally. But a conception of free speech based on democratic deliberation is
nevertheless incomplete. Focusing as it does on the asymmetries of mass media
communication, it does not adequately address the technological changes of the
digital age that make it possible for everyone to participate in electronic
communication, both as speakers and listeners, both as producers and consumers.

Digital technologies foreground certain features of freedom of speech that
were always present but previously remained in the background.

First, speech ranges over the whole of culture, and much of it has only a limited
relationship to deliberation about politics and political issues, the central concern
of democracy-based theories. Many people do talk about politics, to be sure, but
even more people talk about their favorite television show, about their favorite
musical group, about what their child did the other day, about art, popular culture,
fashion, gossip, mores and customs. One can try to insist that all of this speech
really is about politics, but at some point the attempt will become quite strained. It
is far better to acknowledge that speech goes well beyond the boundaries of
deliberation about public issues.

Second, speech on the Internet is interactive. People talk back to each other,
they respond to each other. People are not simply passive consumers of media
products sent to them by broadcast media. They are also active interpreters of what
they find in culture, and they continually exchange their ideas with others.

Third, speech is appropriative: People build on what other people have done. They glom on not only to products of the mass media but also to each other’s work, building new things out of old things. To use the anthropologist’s term, they engage in cultural bricolage. This borrowing and bricolage is non-exclusive appropriation — I take and use what I find, but what I take is still there for other people to take and use too.

Wide ranging creativity that goes beyond politics, interactivity and appropriation are all characteristics of speech on the Internet. But they are characteristics of free speech generally; they were true of speech even before there was an Internet. Digital technologies just make these features of free expression more salient.

The 20th Century concern with speech as a mode of democratic deliberation privileges the delivery of information about issues of public concern to the public, who receive this information through the mass media. Far from denying the importance of that conception, I want to insist that it is only a partial conception, inadequate to deal with the features of speech that the new digital technologies bring to the foreground of our concern. The values behind freedom of speech are about production as much as reception, about creativity as much as deliberation, about the work of ordinary individuals as much as the mass media. Freedom of speech is and must be concerned with the ability of ordinary individuals to create, to produce, to interact, to particulate in culture, to engage in non-exclusive appropriation of ideas and expressions, to make something new out of the cultural materials that lay to hand.

Freedom of speech is not merely the freedom of elites and concentrated economic enterprises to provide media products for passive consumption by docile audiences. Freedom of speech is more than the choice of which media products to consume. It is also expression, creation, production, and interactivity. And freedom of speech in the digital age means giving everyone — not just a small number of people who own dominant modes of mass communication, but ordinary people too — the chance to use these new technologies to participate, to interact, to build, to route around, to glom on, to talk about whatever they want to talk about, whether it be politics, public issues, or popular culture.

Why do I insist that freedom of speech concerns the ability to participate in culture, including in particular popular culture? It is because culture is central to individuality, to being the sort of person that you are. You and I are made out of culture. You talk to people, they talk to you, you anger people, they anger you, you influence people, they influence you, and what gets produced over time is — you! And what that ‘you’ consists in is a moving target. You are always in the process of becoming something other than you already are. Your thoughts are changing, your sense of your values is changing, and they change in large part because you are constantly interacting with people, being influenced by them and influencing them in turn. Participation in culture through the exercise of your freedom of speech is one of the most important and valuable ways through which this cultural
transformation of the self occurs. Interactivity and dialogue, the work of routing around and glomming on, are central to living in the culture that you live in, and to becoming the person that you are. For that reason they are also central to the free speech principle.

A democratic culture does not mean that the state holds elections on what culture should be. Rather it means that everyone gets a fair chance to shape and affect the forms of meaning making that shape who they are and everyone else around them. The ‘democratic’ in the expression ‘democratic culture’ does not mean democratic governance but democratic participation. That participation is important because it allows people to influence each other. But it also has a performative value: When people make new things out of old things, when they produce, when they are creative, they perform their freedom through their participation in culture.

Meiklejohn argued that it is was not important that everyone get to speak, but that everything worth saying was said. A theory of democratic culture, by contrast, argues that what is important is that everyone gets a chance to say it. Freedom arises out of participation in culture, and so the point of a system of freedom of expression is to facilitate this participation. That is why a focus on democratic culture places a greater emphasis than the Meiklejohnian approach on how people appropriate popular culture and make it their own; it values interactivity, creativity and the ability of ordinary people to route around and glom on.

That approach also helps us critique the emerging direction of free speech law. In the United States, at least, I see two disturbing trends in policy debates and court decisions. The first is the increasing assumption that ownership of a distribution network gives the owner a first amendment privilege to control access and the experience of end users. The second is an increasing tendency to see legal protection of intellectual property not as part of a grand bargain that promotes a vibrant public sphere but rather as protection of a property right that owners can use in whatever way they wish and without any thought of how their actions affect the system of freedom of expression.

As I have noted above, this combination of positions in telecommunications policy and intellectual property law promotes the interests of media corporations, who, in their quest for profits, are more likely to treat ordinary individuals as potential consumers of media products, rather than active producers of their cultural world. From that perspective, freedom of expression becomes the freedom to choose which media products to consume. Surely this choice is part of freedom of speech, but it is not everything, or even the most important thing. The new technologies allow individuals, more than ever before, to produce culture rather than passively consume it; they allow people to enact their freedom through cultural participation. But from the standpoint of profit maximisation, this active participation has no independent value except to the extent that it involves the consumption of media goods, in which case it is equivalent to consumption. And to the extent that active cultural participation diverts end users and causes less
consumption of media products, interferes with an expansive definition of intellectual property rights, or challenges corporate technologies of control, it is less valuable than passive consumption; indeed it is positively harmful and must be cabined in.

That is too limited a vision of what freedom of speech is, and what a system of free expression is for. The digital age offers the promise of a truly democratic culture of participation and interactivity. Realising that promise is the challenge of our present era. Technological development produces social conflict, which is fought out in politics and in law. That is how rights change. The point, however, is to make sure that they change in the right way.
Address

Banes of an Income Tax: Legal Fictions, Elections, Hypothetical Determinations, Related Party Debt*
H DAVID ROSENBLOOM†

Abstract

This paper examines the United States tax treatment of so-called ‘related party debt,’ especially in transactions involving entities under common control. As a matter of economics and common sense, it is hard to see how such transactions can meet the normal tests used by United States law to distinguish debt from equity. The hallmark of debt is an unconditional promise to pay a sum certain on demand or at a fixed maturity in the reasonably foreseeable future; a transaction between legal entities under common control each of which is ultimately just a piece of paper is hard to reconcile with this standard. The paper suggests that the widespread recognition of related party debt as debt may stem from unquestioning acceptance of legal fictions like the corporate form, tolerance of elective elements in the tax system, and the common use of hypothetical tests to determine tax consequences.

* This paper was prepared for delivery as the 2003 Ross Parsons Lecture, at the Supreme Court of New South Wales, Sydney, on 28 August 2003. In a slightly different form it was published in Tax Notes International Vol 32, 15 December 2003 at 989. The author would like to thank his colleague Vitaly Timokhov, for his very substantial contribution.

† Director, International Tax Program, New York University School of Law.
My objective in this paper is to bring a little sunshine to the subject of related party debt. Or, to be more specific, I propose to explore the treatment under the income tax of a relationship that purports to be debt entered into between legal entities subject to common control. The income tax treatment — the wide acceptance — of this relationship is puzzling. The untutored eye does not perceive a relationship between separate (if related) persons in any real sense, nor do established authorities and ‘in substance’ analysis support the existence of debt. Yet acceptance of related party debt as debt is generally unquestioned. My thesis is that this ready acceptance flows from certain features of the income tax — legal fictions, elections, hypothetical determinations – which themselves are seldom placed in question, and which combine to make the remarkable phenomenon of related party debt appear perfectly normal and to place it virtually beyond inquiry.

1. Introduction

The income tax system of the United States, like most other income tax systems, generally reflects the Haig-Simons concept of income. This concept holds that 'income is the money value of the net accretion to one's economic power between two points of time.' The definition is important in two ways: it relates income to the accretion, or increase, of economic capacity within a prescribed time period and it points to the subject of taxation — the ‘one’ — whose increase in economic capacity forms the base of the tax.

As an instrument of social and economic policy, the Haig-Simons concept, and thus the income tax, rests on the notion that it is appropriate, reasonable and fair to link the burden that falls on each taxpayer with that taxpayer’s ability to pay. Such ability, in turn, is appropriately, reasonably and fairly determined according to periodic accretions to the taxpayer’s wealth, in the form of income.

1 Tax professionals necessarily remain tethered, to a lesser or greater extent, to the tax systems that are most familiar to them. The discussion that follows is grounded in the income tax system of the United States.


3 There are, of course, differing views on the income tax, and specifically whether it is in fact an appropriate, reasonable and fair type of levy. One view is that a taxpayer’s fiscal burden, his or her contribution to government resources, should be measured not by what flows to the taxpayer but, instead, by what the taxpayer removes from society, or consumes. This approach supports a consumption tax, whether sales, value added, or other form of levy that meters the taxpayer’s burden in accordance with outflows rather than inflows. The approach is bolstered by the notion that accretions to wealth in private hands are requisite to economic growth, and that it is counterproductive to overall welfare to diminish that wealth by taxes on income not spent for consumption.

4 An associated, but independent, rationale for the income tax is that persons with greater ability to pay have evidenced greater utilisation of government services and other benefits that taxes fund. If this were in fact the case, a requirement that such persons pay a greater amount of tax would be tied to the costs those taxes defray.
Since it is generally necessary to expend resources in the pursuit of wealth, the concept of accretion of economic power is logically net of expenses incurred by the taxpayer in order to derive the income subject to tax. Such expenses will vary depending on the type of income-producing activity, and thus the costs reasonably needed to achieve wealth accretion will differ substantially from taxpayer to taxpayer. For example, the costs incurred by a taxpayer to earn income from the manufacture and sale of a product may be greater or less than the costs incurred by another taxpayer in manufacturing and selling a similar product, and may differ even more from the costs incurred by another taxpayer in earning income from the rendering of services. An income tax takes such variances into account by permitting subtractions from the tax base, in the form of deductions, for profit-seeking expenditures, subject to certain limitations. As proper costs of earning income, these deductions are entirely consistent with the Haig-Simons concept.

There are, of course, many differences between the US income tax system (or any other income tax system) and the pure Haig-Simons view of income. Much has been written about the problematic concept of realisation, generally used to mark the point at which accretion is sufficiently objectified for the resulting income to enter into the tax base. In addition, various types of accretion have been declared to fall outside that base, for various reasons. Certain costs of earning income have been inflated, accelerated, or denied deductibility.

In addition to these obvious, and much discussed, divergences from the Haig-Simons concept, there are various features of the US income tax that operate in more subtle ways to subvert the ‘ability to pay’ concept. Among these are three features on which little comment is ever offered but which profoundly affect the economic and legal landscape of the income tax. These are (1) legal fictions, (2) elections, and, (3) hypothetical ‘as if’ determinations — all firmly embedded in the tax system and each worthy of independent analysis. When these features converge, as I believe they do in the relationship known as related party debt, the result is a leading tax reduction tool. Such debt is generally accepted as debt because a tax system that readily accepts legal fictions, countenances numerous elections and relies upon hypothetical tests to determine the tax base does not find it at all strange that related party debt should be treated, in general, just like debt between parties that are not related.

---

5 This is independent of the point that periodic accretion includes, of necessity, periodic consumption.
6 As a general matter, only profit-seeking expenditures should reduce income. The purpose of the limitation is to distinguish between legitimate costs of producing income and expenses for personal consumption. Another major limitation on deductibility involves timing and the factual connection between income and the costs of its production. Thus, a current expense is immediately deductible, whereas expenses conferring a lasting benefit are deducted over a period of time.
7 See, for example, David Shakow, ‘Taxation Without Realization: A Proposal for Accrual Taxation’ (1986) 134 U Penn LR 1111.
2. Legal Concepts, Economic Distortions

If taxable income is the periodic accretion of economic power, taxation must depend on the accurate measurement of that accretion. Clearly, a distortion in any element of the formula (taxable income equals gross income less related deductible costs), whether in the cognisable cost of an asset, recognised gain, or related costs, will result in a corresponding distortion of the base subject to tax. As distortions multiply, the gap between accretion of economic power and taxable net income widens. Distortions are, of course, exacerbated by unsurprising taxpayer efforts to bend elements of the measurement process to the goal of minimising liability.

Intentional tax reduction involves, in essence, only a few basic techniques — reduction of gross income, inflation of deductible expenses, deflection of the tax base to others standing in a favoured tax position. All these techniques are greatly enabled by features of an income tax system that lend themselves to manipulation at the taxpayer’s will.

A. Legal Fictions

Most nations that follow a rule of law accept as given certain important legal fictions. Indeed, a basic aspect of legal training, absorbed by law students throughout the world, is that the ‘law’ operates with unshakable acceptance of such fictions. In these circumstances, it would be practically impossible for any jurisdiction to disregard fictions altogether in shaping its rules of taxation, if only because those fictions have important non-tax consequences.

One ubiquitous legal fiction is the corporate form — as fictitious a construct as could be imagined. ‘Legal persons’ owing their existence to rules of law, corporations generally can act in their own capacity, sue and be sued, take and hold property. They were adopted in 17th Century England as a mechanism to spur trade by shielding merchants from unlimited liability, and have always been an economic extension of their owners enjoying a special privilege granted by the sovereign.8 Developments in legal and economic systems in most industrialised countries and the advent of public shareholding (whereby ownership by large numbers of people became common) gave the corporation an entirely new role. In addition to being a business accommodation, a security mechanism, the corporation became a virtually autonomous economic unit conducting business without control, or even oversight, by its owners.9 The corporation thus came to form associations on its own, and to function itself as a shareholder and partner in business associations. In this manner, the corporation has become so accepted, so

8 US income tax law nevertheless recognises the economic identity between corporation and shareholder in various places where the recognition is beneficial to taxpayers, such as S corporations, the deemed paid foreign tax credit and the treatment of limited liability companies. See IRC (Internal Revenue Code) ss1341, 902, 960; 26 CFR (Code of Federal Regulations) ss301.7701–2, 3.

9 Thus, many corporations operate, in reality, wholly independently of their shareholders. It is clear that these owners, with some exceptions, have no say in regard to day-to-day operations of these large constructs. Managers (who need not be owners) are basically free of owner-determined constraints, except in extreme cases.
anthropomorphised, that it is perceived not only by those trained in the law, but in the popular press and by the public generally as having a real existence apart from both its owners and its managers.\footnote{10}

The corporate ‘person’ is also malleable. Unlike the human being with which the tax law equates it, the corporation is perpetual; but it may be quietly terminated at the will of managers and shareholders. The corporation can also be brought quickly into life and then transformed into a different corporation, loaded with features to support ‘economic substance,’ or merged into other corporations by filing appropriate documents with appropriate authorities. It cannot be deprived of life and freedom under the US Constitution.\footnote{11} In the modern world, it can migrate to countries other than the one in which it was formed,\footnote{12} facing no immigration issues, with fiscal authorities in the new motherland welcoming the new corporate citizen without extensive inquiries into its lineage, substance, or purpose. In brief, the corporation has much greater possibilities than the natural person — requiring the tax laws to adopt commensurately multi-faceted rules and broadly (if quietly) affecting tax policy at both the stage of its formulation and that of its implementation.

Corporations, however, are only pieces of paper. There is nothing ‘real’ about a corporation, nothing tangible, nothing capable of acting except through human agents, nothing with an independent will, intent, or moral capacity of its own. It is hard to understand what prosecutors mean or intend when they occasionally speak of criminal penalties for corporate malfeasors.\footnote{13} Penalties that do not affect individuals have little deterrent power, and the individuals harmed by corporate prosecutions may easily be persons (owners) other than the malfeasors (managers).

Non-legal policymakers in the field of taxation, especially economists, are perhaps less likely to think about entities, including corporations, as independent persons, as opposed to extensions of the persons who own or control them. And it may be that a truer concept of the modern corporation is an economic business

\footnote{10} The first indication of an abusive use of the corporate form in US tax matters came in the 1930s when, as a result of differences in tax rates, corporations were used to park income away from the higher rates imposed on individuals. In 1934, Congress introduced the personal holding company rules, imposing a penalty tax on these ‘incorporated pocketbooks’ by taxing earnings accumulated in corporate solution. The sanctity of the corporation was not questioned — although a simpler and economically more acceptable result would have been achieved by ‘piercing the corporate veil’ and attributing income of the incorporated pocketbook to its owners.

\footnote{11} With few exceptions, corporations have all the constitutional and legal rights of a natural person. Although corporations cannot vote, they can lobby elected members of the government and cannot be deprived of life, liberty, or property without due process of law. \textit{First National Bank of Boston v Bellotti} 435 US 765 (1978).

\footnote{12} See, for example, IRS Field Service Advice Memorandum 200117019 (27 April 2001), involving a corporate ‘continuation’ from the United States into Canada.

\footnote{13} See, for example, Larry Thompson, Deputy Attorney General of the United States, ‘“Zero Tolerance” for Corporate Fraud’ \textit{Wall Street Journal} (21 July 2003) at A10: ‘It is a bedrock principle of American law that business organizations, including corporations, may be held to account by the criminal law for the wrongdoing of employees or agents.’
unit, and that this concept should replace, at least for fiscal purposes, acceptance of the corporate form as a legal person. Certainly US law relating to international taxation, with its rich blossoming of complex rules, sub-rules and interpretations of rules, could be vastly simplified by looking through the form of the controlled foreign corporation (and perhaps other corporations) to the substance of its controlling shareholders and developing rules accordingly. Nor are the complications limited to the international tax field. The economic justification for regarding a group of ten corporations as ten separate persons is hardly compelling. There may well exist non-tax justifications for creation of the group, but it is not foreordained that the tax laws should respect the results, particularly when the multiplicity of choice may be damaging to the goals of those laws. By assigning tax consequences to corporations (and other entities), a tax system effectively permits choices of tax results unattended by economic consequences. The legal fiction accords taxpayers a substantial measure of control over the computation, and therefore ultimately the amount, of tax they are called upon to pay.

Since a corporation is a convenient mechanism for collection of tax, the US system considers it an income taxpayer and imposes tax on it. The corporation is thus a non-transparent person for US tax purposes, with its own tax base. Whether a corporate tax could be made consistent with otherwise ignoring the corporation as a separate person is an interesting and intricate question. At a high level of generality, there does not appear to be a fundamental inconsistency between collecting tax from the corporation and according it something less than full, individual-equivalent, ‘person-hood.’

In any event, other entities, constructed pursuant to other legal fictions, are generally regarded as transparent by the tax laws, so that income and tax consequences pass through them to their owners. Once the corporation is accepted for tax purposes as an independent person, it must be differentiated from these other persons that give rise to dramatically different tax consequences. This is not easy, and countries have taken varying approaches to the task.

**B. Elections**

Elections, like legal fictions, are probably inevitable in an income tax system. In numerous instances, a choice between alternatives would be unjust, or would lead to otherwise undesirable results. In addition, elections are politically attractive: they satisfy everyone. Forced to choose between solution A and solution B, the legislator passes the choice to the taxpayer, who can be counted upon to make the choice in his or her own favor. Everyone, ostensibly, wins.


15 The US tax world is made more complicated by the long-standing existence of transparent corporations and non-transparent trusts and estates. See IRC ss1301 and following (S corporations); ss641 and following (trusts and estates).

The problem is that, like legal fictions, elections have costs. They add complexity to a tax system, because an explicit election requires rules to govern each of the elective possibilities. In addition, elections affect the fairness, or perceived fairness, of the system when their availability is limited to certain classes of taxpayers, or when they can be revoked from time to time, as a taxpayer’s underlying situation, and therefore his or her tax objective, changes. The overall impression tends to be a tax system subject to the taxpayer’s control, a concept antithetical not only to an income tax but to any tax. Taxation is compulsory, not elective, and involves no obligation on the part of government other than the provision of public goods and services. Taxation is thus different from raising government funds through charitable contributions, issuing debt or establishing slot machines. An election for a taxpayer — or certain taxpayers — to pay tax at either a 15 per cent rate or a 20 per cent rate would be absurd.

Some elections in the income tax are explicit: the taxpayer is allowed to choose whether to allocate expenses according to the cost basis of assets or their fair market value; interest expense of a foreign person is attributed to a host country business using either of several alternative methods. In these instances, the policymaker might ask whether it is really necessary to provide such alternatives. Why would a blended approach not work for everyone? Of course, political forces push in the direction of electivity (they always do — elections are never to the taxpayer’s detriment), but such forces also push in the direction of abandoning taxation altogether. The policymaker will strive for compromise and, in many instances, will find it impossible to withstand the temptation to make everyone content. But the use of elective alternatives undeniably contributes to the vulnerability of the tax system on a variety of grounds.

Nor are elections necessarily explicit — and therefore readily identifiable to persons attempting to evaluate them. When taxpayers have a choice of transactions that are economic alternatives but that produce different tax consequences, the choice raises as many issues as, and perhaps even more than, the explicit election. Here the problem is not the multiplicity of rules. Different rules are pre-existing, because the system has identified a difference in transactions: payment of a dividend is, arguably, different from payment on a derivative contract. When, however, a taxpayer may use either transaction to obtain identical, or very similar, economic results, most of the issues relevant to explicit elections are present. The economic difference between a dividend and payment on an equity swap is arguably not compelling. There will be differences in the creditworthiness of the payors, but those differences may not justify different tax consequences.

---

17  26 CFR s1.861–9.
The two types of elections — explicit and transactional — and the relationship between them are nicely illustrated by the US experience with rules of entity classification and their incarnation in the famous ‘check the box’ regulations of recent vintage. Under the pre-existing *Morrissey* factors, which stemmed from a Supreme Court decision in 1935, entities were classified as corporations, partnerships, or trusts according to a mechanical application of factors: an objective of carrying on business and dividing the profits, associates, limited liability, centralisation of management, free transferability of interests, continuity of life. The original regulations adopting these factors after *Morrissey* contained a bias against the association (taxable as a corporation). There were an even number of factors to weigh and corporate status was assigned only when a majority of the factors pointed to association status. Tax administrators were concerned that professionals such as doctors and dentists would incorporate their practices to avoid application of what were then extraordinarily high rates of individual taxation. But as often happens in a complex world, the regulations remained on the books after the issues putting pressure on the question of entity classification changed. In the late 1960s and 1970s, the focus became tax shelters that depended on an entity affording limited liability to investors but that was transparent, allowing a flow-through of losses and credits. The bias against the non-transparent corporate form came to operate to taxpayers’ benefit, as they took advantage of the regulations to engage in a great deal of shelter activity.

That activity put increased pressure on the precise meaning of the various *Morrissey* factors, each of which developed a high gloss and a nuanced meaning. In addition, with expanding internationalisation in the 1980s, the entity classification rules came to apply with frequency to exotic foreign entities. It was far from clear how to apply the factors, with their refined interpretations, to these novel constructs. Generally, taxpayers could achieve the results they wished for a foreign entity, transparent corporation or non-transparent partnership, with little sacrifice of economic substance, by tweaking the entity to control the result under the *Morrissey* tests. The Internal Revenue Service was finding it necessary, in order merely to keep up, to track the latest developments in regard to Peruvian SAs and entity offerings in Niger. And the stakes rose substantially when states of the United States adopted, one after another, the now familiar limited liability companies which could, in general, be used instead of the corporate form to limit

---

20. 26 CFR s301.7701(a)–2. The regulations were adopted by TD (Treasury Decision) 8697 (18 December 1996), 1997–1 Cumulative Bulletin 215, and went into effect generally as of 1 January 1997.


liability without meeting the *Morrissey* tests for an association. This permitted foreign persons investing in the United States to engage in many of the same elective transactions as US persons using foreign entities (or US limited liability companies) to invest abroad.

Thus it was that the Service raised the white flag of ‘check the box,’ allowing taxpayers an explicit election in place of the transactional election they had previously enjoyed. It was believed that costs in terms of tax administration would be reduced, the situation would be more open and accessible to the uninformed, and the law would be, in some sense, more ‘honest.’ This was a forthright and proper action by the Service, *assuming there had to be any election at all.* Years earlier, the Treasury had proposed that there ought not to be an election here — that if an entity afforded limited liability to its owners, that fact in and of itself should result in corporate classification. Protests from taxpayers (citing *Morrissey*) led the Service to withdraw the proposal and continue with the transactional election based on the *Morrissey* factors. It was only when the government’s patience with the factors ran out, and frustration with the implications for tax administration began to grow, that the transactional election gave way to the explicit election of ‘check the box’.

Many lessons for policymakers (and students of taxation) lurk in this story. If the legal fiction of the corporation as a separate non-transparent person is respected for tax purposes, there must be at least two sets of rules, one for the corporation and one for other entities, such as partnerships, that are transparent. But there is no reason why the dichotomy cannot turn on a difference of substance, as was suggested in the 1970s. The gift of an election permits taxpayers to have it as they wish, transparent or non-transparent, at will. Moreover, since the formation and dissolution of entities is simple, migration from one status to the other is not a problem. Allowing taxpayers to choose whether the legal fiction through which

24 See, for example, IRS Revenue Ruling 93–6, 1993–1 Cumulative Bulletin 229 (Colorado); IRS Revenue Ruling 93–38, 1993–1 Cumulative Bulletin 233 (Delaware); IRS Revenue Ruling 94–51, 1994–2 Cumulative Bulletin 407 (New Jersey).
25 For the reasons that led Treasury to abandon its initial approach to entity classification, see Notice of Proposed Rulemaking, Simplification of Entity Classification Rules, 61 FR 21989 (13 May 1996).
26 An independent question addressed by the ‘check the box’ rules — whether a transparent entity owned by a single person should be disregarded — represents a distinct, and arguably more problematic, tax policy judgment. The rule for so-called ‘single-member entities’ was the product of rigorous logic (how could a transparent entity owned by a single person have an independent existence for tax purposes?) but, it would appear, little independent analysis.
27 Notice of Proposed Rulemaking, Classification of Limited Liability Companies (17 November 1980), Proposed Regulation s301.7701–2(a)(2) (1980): ‘... an organization will be classified as an association [taxable as a non-transparent corporation] if under local law no member of the organization is personally liable for debts of the organization.’
28 The withdrawal was accomplished by a Withdrawal of Notice of Proposed Rulemaking, Classification of Limited Liability Companies, 48 FR 14389 (4 April 1983). The decision in *Morrissey* was, of course, not constitutionally mandated.
29 Trusts tend to be a stepchild in this analysis, though the many advisors who make a living by sweeping the Internal Revenue Code for gold are well aware that the trust form exists, and can be exploited.
they operate should be transparent or non-transparent seems both unnecessary as an intellectual matter and questionable as a matter of tax policy.

The US income tax allows many explicit and transactional elections — situations where taxpayers are permitted to choose one tax treatment or another without substantive cost. It would be foolhardy to venture the judgment that such elections should be abolished entirely, but policymakers would do well to recognise that elections burden the system, in terms of fairness, appearance (of fairness) and complexity. The Internal Revenue Code did not come to look the way it does by simply characterising for tax purposes the persons and transactions that could be foreseen to engage in economic activity; the heft of the document is due in large part to an understandable taxpayer desire for choice, coupled with an equally understandable politician desire for accommodation. At the end of the day, however, a choice for everyone results in a separate tax law for everyone, and that would not be a desirable policy option for a variety of reasons.

To make the point another way, if ‘check the box’ represents a sound response to the problem of entity classification, why not employ the same technique for the equally difficult issue of instrument classification? Taxpayers and tax administrators commonly encounter difficulties in distinguishing debt from equity. Could the answer be to permit the taxpayer to ‘check the box’? For that matter, if ‘check the box’ is suitable for entity classification in the United States, presumably it would be equally suitable for France or Japan. ‘Check the box’ has produced some serious problems for the US income tax, but either of the foregoing perfectly logical suggestions would render matters worse.

Aggressive use of the entity classification rules and the corporate form allows sophisticated taxpayers to obtain advantageous tax treatment. This has both policy and practical implications. A tax system will not be considered fair if some taxpayers may restructure their forms of doing business, without economic change, to achieve results that are unavailable to other taxpayers. Furthermore, there are obvious costs to such restructuring in the form of reduced tax revenues and the burden placed on tax administration.

Since an election, by definition, results in a reduction of income tax liability (without a reduction in economic accretion), it can be viewed as a tax expenditure.30 Although the government periodically estimates revenue losses attributable to tax expenditures, no study estimating the fiscal effect of transactional elections seems ever to have been undertaken. Nor, as a matter of methodology, is it clear how measurement of the effect of taxpayer changes of forms and characteristics to attain tax reductions could be accomplished.

From multiple points of view, elective provisions of the Code tend to operate at cross-purposes with the goals of the income tax. Elections, express and

30 Tax expenditures are defined under section 3(3) of the Congressional Budget and Impoundment Control Act of 1974, Public Law Number 93-344 88 US Statutes-at-Large 297, as ‘revenue losses attributable to provisions of the Federal tax laws which allow a special exclusion, exemption, or deduction from gross income or which provide a special credit, a preferential rate of tax, or a deferral of tax liability.’
transactional, add complexity and revenue cost, and burden the effective administration of the tax laws. Their use almost certainly cannot be eliminated, but a degree of sensitivity to the harm they cause would be highly desirable.

C. Hypothetical Determinations

Like legal fictions and elections, hypothetical tests are needed in the income tax: transaction A to be analysed as if it were transaction B and the result applied to transaction A. And this is doubtless the case without regard to legal fictions. It is not hard to imagine situations in which transactions having no fictional aspects really must be subjected to an ‘as if’ analysis: a father sells widgets to his son and the tax law asks what the price of the widgets would have been if the transaction had been between parties that were not related. It is clear, however, that legal fictions multiply both the number and variety of situations in which transactions cannot be accepted at face value, and that such fictions have thus contributed greatly to the proliferation of hypothetical inquiries that must be undertaken if the tax system is to remain compulsory.

The poster child for hypothetical determinations is, of course, the ‘arm’s-length’ method, a technique employed by the income tax to determine and test pricing for transactions between related persons. There are several well-documented problems with this method. The standard employed is imprecise, because normally there is no such thing as a single ‘arm’s-length’ price for transactions between unrelated parties, since the price actually used will depend on a nearly infinite number of variables relating to the knowledge, understanding, needs and skills of the parties. In addition, the data required to apply the standard are hard to find in any form, and second- and third-best substitutes for ideal data are easily manipulated by persons with an understanding of the rules of the game. The promised impartiality of the method thus dissipates into a familiar trading of horses in light of the relative leverage of the parties — not the leverage of the (related) parties to the transaction being tested, but the leverage between taxpayer and tax administrator when a controversy surfaces. In the United States, at least, the taxpayer can threaten to extend the examination, using a variety of procedures and means at his or her disposal; the administrator can threaten penalties and rely on the expense of litigation. Some things never change.

31 See IRC s482 and the regulations thereunder. The purpose of section 482 and the regulations is to place ‘a controlled taxpayer on a tax parity with an uncontrolled taxpayer by determining the true taxable income of the controlled taxpayer.’ 26 CFR s1.482–1(a)(1). Or, in the somewhat starker terms used by the Tax Court in Your Host v Commissioner, section 482 ‘is designed to remedy only one abuse: the shifting of income from one commonly controlled entity to another.’ 58 TC 10 (1972), 24, affirmed, 489 F 2d 957 (1973). In determining the true taxable income of a controlled taxpayer, ‘the standard to be applied in every case is that of a taxpayer dealing at arm’s length with an uncontrolled taxpayer.’ 26 CFR s1.482–1(b)(1).

32 See, for example, Stanley Langbein, ‘The Unitary Method and the Myth of Arm’s Length’ (1986) 30 Tax Notes 625.
The difficulty with the ‘arm’s-length’ method that is most troubling, however, is neither its falsely promised precision nor the difficulties of obtaining the information needed to employ it, but its very nature: a hypothetical test. A transaction between related parties (and there are a potentially infinite number of related parties, once the corporation is accepted as an independent person) is analysed as if the parties were not related. Or, as I suggested at a conference in the 1980s, the question asked by the ‘arm’s-length’ method is whether, if you had a brother, he would like cheese.33

There is, of course, a reason why related parties that engage in a transaction are related, and that reason, however described, is almost certainly lacking in the case of parties that are not related. Moreover, the reason almost certainly has an important influence on any transaction in which the parties engage. Thus, the attempt to ask what the parties would have done, how they would have dealt with one another, if they had not been related (as the ‘arm’s-length’ method effectively asks) is, in the most fundamental sense, nonsense; it is doomed to send the asker wandering into speculative considerations.34 Furthermore, the question is inherently ambiguous. What aspects of the tested (related party) transaction are to be accepted before the ‘as if’ test is applied — everything save the actual price? Suppose it can reasonably be concluded that, if the parties had not been related, they would not have engaged in a transaction at all, or would have engaged in a different kind of transaction, or would have satisfied their business needs with other persons in other ways? What does ‘arm’s-length’ — what does ‘as if’ — imply? The level at which the inquiry is pegged is not specified, leaving the analyst to determine that crucial point for himself. As a result, tax consequences flow from a tug of speculations between taxpayer and tax administration, inevitably giving rise to spirited controversy (and a dispersion of results across the taxing jurisdiction).

The point here is not to weigh the ‘arm’s-length’ method against potential alternatives, but to note the problems that the method, a hypothetical test, inevitably spawns. Such ‘as if’ testing and determining of tax results are not uncommon in the US income tax and, by reason of the prominent leadership role that the United States plays in the field of taxation, in other income tax systems as well. As most lawyers would acknowledge, nothing is ever just like anything else.35 ‘As if’ tests invariably produce unsatisfying results because they veer into the subjective and therefore the manipulable. Once again, a tax world in which such tests are abandoned altogether is probably not possible, but there is no good reason to employ them except when they are truly necessary.

33 It is no accident that the name of the taxpayer in the ‘cheese examples,’ Examples (2), (3), and (4) of 26 CFR s1.482–4(f)(3)(iv), is Fromage Frere. Alas, newly proposed regulations would dispense with these examples. See Notice of Proposed Rulemaking, Treatment of Services under Section 482, Proposed Treasury Regulation s1.482–4(f)(3) 68 FR 53447.
34 See Staff of Joint Committee on Taxation, General Explanation of the Tax Reform Act of 1986 (Washington: JCP, 1987) at 1014: ‘Observers have noted that multinational companies operate as an economic unit, and not ‘as if’ they were unrelated to their foreign subsidiaries.’
35 The observation is not intended to conflict with the classroom observation of the late Paul Freund, Professor of Constitutional Law at Harvard Law School, that, on the other hand, everything is just like everything else.
3. **Related Party Debt**

One cost that most income tax systems allow as recoverable, or deductible, in most circumstances is business-related interest. This is simply the cost of money borrowed, the functional equivalent of rent paid for a machine.\(^{36}\) The item obtained for this cost — temporary use of money or machine — is used to earn income that will go first to recover the cost of the temporary use and then, it is hoped by the borrower or renter, return a profit. This profit, of course, becomes the base of the tax. It is neither extraordinary nor objectionable, as a general matter, that the cost of funds should be so viewed in the context of an income tax. Debt and interest are normal features of engaging in an income-producing business.

Given the purposes of the income tax, and the rationale for the interest deduction, related party debt is an oddity. As its name indicates, this is indebtedness between parties having a relationship to one another. There are, obviously, many ways of defining that relationship — of establishing the line that separates related parties from parties that are not related — but these definitional matters are beside the present point. Regardless of where the line is drawn, in most jurisdictions the line is of only limited, and specific, significance.\(^{37}\) As a general matter, related party debt is seen as debt, and in that capacity viewed as giving rise to income from money lent in the hands of the creditor and deductible expense for the debtor.\(^{38}\)

Related party debt is a principal tool of the tax planner. Since related parties may differ between themselves in any number of tax characteristics — residence, method of accounting, functional currency, exempt versus non-exempt status — the interest deduction and commensurate income from related party debt may produce overall beneficial tax results within a single economic unit. This may be of little consequence when individuals (father and son, for example) are the parties that are related. But related party debt is not limited to, or even most common in, such circumstances. Since the tax laws accept legal fictions and view each legal entity as a separate person, related party debt is found in the greatest volume between commonly controlled entities. In these circumstances, the differences in taxation between debtor and creditor can be used to obtain a tax-free return from a source country, drive down effective rates of tax, achieve a deduction without a commensurate income inclusion, and in many, many other situations.

---

\(^{36}\) See *Deputy v du Pont* 308 US 488 (1940), 498, where the Supreme Court authoritatively defined interest as ‘compensation for the use or forbearance of money.’ In a later decision, *Dickman v Commissioner* 465 US 330 (1984), 339, the court adopted a more ‘economic’ approach, defining interest as the equivalent of ‘rent’ for the use of funds.

\(^{37}\) See, in the United States, ss163(e)(3), 163(j) and 267 of the Internal Revenue Code for examples of rules that limit the interest deduction on related party debt, and s904(d)(3)(C) for a different type of special treatment.

\(^{38}\) See IRC ss61(a), 163.
Debt can also easily be hybridised, or treated as equity in one tax jurisdiction while retaining its status as debt in another, and related party debt is a critical feature of such instrument arbitrage. And entity arbitrage, of the sort commonly referred to as the ‘hybrid branch,’ would have far less utility without the use of related party debt. In such cases, one tax jurisdiction perceives a non-transparent entity (outside the jurisdiction) while the other sees a transparent, and disregarded, entity. Related party debt owed to the hybrid supports interest that will drive down the tax base of the jurisdiction from which the payment is made, while not being recognised in the jurisdiction that deems the entity, and thus the interest payment, not to exist at all.

In fact, when an observer casts around for related party debt, it is everywhere — transferring the incidence of taxation away from one taxpayer and into the hands of another, lower-taxed or non-taxed, often foreign, related taxpayer. This versatile technique is accepted generally as the equivalent of debt between unrelated parties — subject, of course, to policing for an ‘arm’s-length’ rate of interest and perhaps extra scrutiny in suspected thin capitalisation situations, but never otherwise in substantial doubt.

There seems to be only one serious problem with related party debt: by most standards of economics, ‘substance,’ or common sense, it is not debt. That is, related party debt is generally not compensation for money lent by one person to another. Rather, it is a transfer of funds from one incorporated pocket to another, usually for tax-reduction purposes. Only a tax professional, considering indebtedness between commonly controlled entities, would perceive a similarity

39 A modicum of care is, however, needed. In Laidlaw Transportation v Commissioner 75 TCM 2598 (1998), failure to observe the original terms of a purported debt instrument (treated as equity by a Canadian parent corporation) caused the Tax Court to recharacterise the instrument as equity for US tax purposes and disallow interest deductions to US subsidiaries.

40 IRS Notice 98–11, 1998–1 Cumulative Bulletin 433, an early warning that such transactions might be questionable on policy grounds, has had no effect.

41 Related party debt is not unknown in a purely domestic context, but the rules relating to consolidated returns for affiliated groups deprive it of much of its utility there.

42 An exception would be the Tax Court Memorandum decision in Laidlaw, above n39, where the court purported to look to ‘substance’ in classifying ostensibly related party hybrid debt as equity for US purposes and denying interest deductions claimed by the taxpayer. In Laidlaw, however, original maturity dates had been extended and interest payments had been re-lent to the putative borrower, and the court, at 2624 emphasised these facts: ‘The substance of the transactions is revealed in the lack of ‘arm’s-length’ dealing between [a foreign corporation] and petitioners [US subsidiaries], the circular flow of funds, and the conduct of the parties by changing the terms of the agreements when needed to avoid deadlines. The Laidlaw entities’ core management group designed and implemented this elaborate system to create the appearance that petitioners were paying interest, while in substance they were not.’ Thus, the court was able to sidestep the broader implications of its earlier observation that possibilities, not actual events, should drive substantive analysis, at 2616: ‘if a transaction is controlled by related entities, the form and labels used may not signify much because the parties can mold the transaction to their will’, citing Anchor National Life Insurance v Commissioner 93 TC 382 (1989), 407. Laidlaw is a Memorandum decision, with no precedential value.
to raising funds from unrelated parties. As noted, interest is generally deductible in the unrelated party context because it is a legitimate cost of the funds needed to operate a business. The application of this rationale in a related party context, where no funds are being raised, is one of the tax miracles of our time.

The US Tax Court came close to this point in an entirely different context in *Albertson’s, Inc v Commissioner*, which involved the deductibility of ‘interest’ accrued on certain unfunded deferred compensation arrangements. The court noted that ‘one cannot lend that which one never had a right to possess, and petitioner [seeking to deduct interest expense] cannot borrow money that it always had the right to possess.’ [Emphasis in original.] By the court’s simple reasoning, one cannot borrow from oneself. When an entity ‘borrows’ from a related entity, the resulting use of funds is not much different, in substance or reality, from the use asserted in *Albertson’s*. But for the legal fiction that recognises each entity as independent, borrower and lender are one and the same.

Related party debt appears to reflect an admixture of techniques, assumptions and methods of approach to tax policy that have become ingrained in the income tax laws of most nations, and certainly in those of the United States. It builds on the legal fiction of the corporation, relies on the common use of hypothetical tests to determine tax consequences and survives in a policy environment that readily countenances elective elements of taxation. As noted previously, each of these features of the tax system has side effects that, independently, produce questionable, sometimes ridiculous, results. Together, they combine to support a proliferation of related party debt.

Under the US Internal Revenue Code, the deductibility of interest on debt represents a sharp distinction from the treatment of equity, dividend payments on which are generally non-deductible. The distinctive tax treatment of each of two common ways of financing a business reflects, in essence, a perceived difference in the nature of the relationship between the party providing funds and the party raising funds for operation of a business. Though both interest and dividends may be viewed, in non-tax parlance, as costs of raising capital, there is a marked difference, without regard to tax laws, between equity and debt. A holder of equity is an owner of the business and dividends may fairly be viewed as a return for the risk of investment. A debt-holder, on the other hand, is not an owner of the business but, rather, a renter-out of funds, entitled to demand ‘a fixed sum at some period set in advance.’ The risk assumed by the debt-holder is limited, and the

---

43 *95 TC 415* (1990).
44 Id at 423.
45 It is true that, since the corporate fiction is accepted for non-tax as well as tax purposes, there can be real non-tax consequences, for example in a case of insolvency, to related party debt. But there is no reason why the mere possibility of non-tax consequences should make a general case for equating related party debt with unrelated party debt.
cost of raising funds through an issue of debt is much closer in nature to other costs incurred in a business venture.\textsuperscript{47}

Dividends are paid to the equity-holder out of the profits of the business. A lender, in contrast, receives income from a loan even before, and regardless of whether, the borrower is profitable. For that reason, debt and equity are not merely different in kind but, to some extent, operate at cross purposes. There will not be any profit for the equity holder to enjoy if costs, including interest costs, are too large.

Policymakers and courts in the United States have struggled for years with the problem of distinguishing debt from equity in particular cases. Often the inquiry has focused on whether an identified payment represents deductible interest or a non-deductible dividend. The US Congress, in enacting section 385 of the Code in 1969, enjoined the Treasury to look into this subject and develop suitable rules to distinguish debt from equity.\textsuperscript{48} Treasury attempted to do this over a period of years in the late 1970s and early 1980s before giving up, withdrawing its ‘final’ regulations on the subject and sending it back for resolution on a case-by-case basis in the courts.\textsuperscript{49} The judiciary had already developed 13 or more ‘factors’ that would permit the testing of an instrument to determine how to categorise it.\textsuperscript{50} At first glance, such a multifarious means of reaching a determination appears to reside in the eye of the beholder, and an objective commentator could be forgiven for concluding that, in a given case, the factors can usually be massaged to steer the decision-maker in the desired direction.

In 1994, the Internal Revenue Service issued a Notice providing (in an all too common but, alas, none too helpful, formula) that the characterisation of an instrument as debt or equity would depend on the instrument’s terms and all surrounding facts and circumstances.\textsuperscript{51} Following the approach of section 385, the Notice went on to enumerate eight factors bearing on the determination, with an admonition that the weight to be accorded any of them would depend on all the facts and circumstances.

Interestingly, the first three of the eight factors enumerated in the Notice are: (1) whether the instrument contains an unconditional promise by the issuer to pay a sum certain on demand or at a fixed maturity in the reasonably foreseeable future; (2) whether holders of the instrument possess the right to enforce payment

\textsuperscript{47} See, for example, \textit{In re Lane} 742 F 2d 1311 (1984). The discussion in the text is obviously not intended to deal comprehensively with the difference between debt and equity as a theoretical matter. It could, for example, be maintained that there should be no deduction for interest on debt. But the dichotomy described in the text is found in most tax systems and presumably reflects a common judgment that distinctions of the sort outlined in the text are justifiable.

\textsuperscript{48} Public Law Number 91–172, s415(a). Notably, the section applies only to corporations, leaving case law as the only authority for transactions involving other entities.


\textsuperscript{50} A survey of Tax Court decisions over the period from 1955 to 1987 actually identified 26 separate factors that went into the analysis. Paul Robertson, Zoel Daughtrey & Daryl Burckel, ‘Debt or Equity? An Empirical Analysis of Tax Court Classification During the Period 1955–1987’ (1990) 47 Tax Notes 707.

\textsuperscript{51} Notice 94–47, 1994–1 Cumulative Bulletin 357.
of principal and interest, and (3) whether the rights of holders are subordinate to the rights of general creditors. Although the Notice prophylactically states that no particular factor is conclusive in making a determination, it is hard not to ascribe some significance to the fact that the ‘unconditional promise’ factor appears in the list well in advance of the ‘label’ factor. The Notice, in that regard, is in step with much of the sprawling case law, which is apt to conclude that the real issue — what a Technical Advice Memorandum issued by the Revenue Service in 1998 called ‘the important issue’ — is whether there was a ‘genuine intention to create a debt, with a reasonable expectation of repayment, and [whether] that intention comport[ed] with the economic reality of creating a debtor-creditor relationship?’ This accords with the observation, for example, that debt provides for ‘repayment absolutely and in all events, or that principal be secured in some way as distinguished from being put in hazard.’

The debt/equity issue goes off in many, sometimes subtle, directions, and there can be circumstances in which an apparent entitlement to repayment may be accompanied by terms that place the holder at sufficient risk to generate a strong flavor of equity. It seems clear, however, that a fixed right to repayment, on the one hand, and participation in the risk of success of the business, on the other, are very basically at odds. In the absence of an explicit or clear participation in risk, the ability of a holder to demand repayment from a debtor at a reasonably near-term date certain is a definite hallmark of debt.

Many difficulties encountered in distinguishing debt from equity affect instruments issued by one related party to another. US courts purport to take this circumstance into account in their deliberations, if only obliquely, but there is little evidence that they have ever struggled with the implications of the relationship. Certainly, there is no recognition of the possibility that there might be a difference between related parties that are human beings and related parties that are entities under common control.

Related parties of the latter type warrant close scrutiny in light of the principles that have been developed in the debt/equity wars. As noted earlier, such related parties are potentially infinite in number. Entities constituting independent parties can be formed and unformed at will in the number that the person or persons

---

52 The other listed factors are whether the instrument gives holders a right to participate in management of the issuer; whether the issuer is fully capitalised; whether there is identity between the holders of the instrument and the owners (stockholders) of the issuer; the label placed upon the instrument, and whether the instrument is intended to be treated as debt or equity for non-tax purposes. See also IRS Revenue Ruling 2003–97, 2003–34 IRB 380, adding ‘the intent of the parties,’ which would seem difficult to establish independently of other factors.

53 IRS Technical Advice Memorandum 199910046 (16 November 1998).


55 Hartman v Commissioner 17 TCM 1020 (1958), 1023.

56 Notice 94–47, above n51, lists as a factor the identity between holders of an instrument and shareholders.

57 Again, the decision in Laidlaw, above n39, provides a hint of an exception. See above n42.
ultimately in control desire. Whatever else may be said about acceptance of each entity as separate from the others, in the context of an ostensible ‘debt’ owed by one such party to another, it strains credulity to say that ‘repayment absolutely and in all events’ is envisioned or provided for. Two pieces of paper owned ultimately by a single economic interest cannot do business together, no matter what legal sophists say. There is lacking in such ‘transactions’ the frictional element that permits an outsider, such as tax authorities, to give credence and respect to the results.58

It follows that the tax law must resort to its customary fallback, the hypothetical test, to determine whether, if the entities had not been related, they would have transacted as they ostensibly did. That type of inquiry, in the case of related party debt, focuses on the interest rate charged — whether it meets the ‘arm’s-length’ test for rate, terms and so on — but perhaps also on the fundamental question of whether the transaction would have occurred at all if friction had been present. That, of course, is an unanswerable question in a related party debt/equity context: not difficult, not complex, unanswerable.

Nevertheless, US law has proceeded to answer it, using the 13 factors (and others) heretofore described, or whatever substitute passes for a debt/equity test in the jurisdiction with decision-making authority. If the transaction passes muster under the test for thin capitalisation — that is, if the planners of the transaction have successfully drafted the instrument to accord with the expected jurisprudence — a finding of debt is entirely possible.

This approach makes related party debt a transactional election of the type described earlier, which the tax planner may choose, or not choose, to put in place. It would seem that this fact, by itself, would be inconsistent with leading indicators of debt status under the tests normally applied. For whatever an instrument says on its face, whatever artfully drafted surrounding documents may say, an obligation owed by one piece of paper to another (both controlled by the same interests) is hard to classify as an ‘unconditional promise’ to pay a ‘sum certain’ at a ‘fixed maturity date’ (or on demand) within the ‘reasonably foreseeable future.’ It may have the appearance of touching each of these bases, but the reality of the situation belies satisfaction of any of these points. There is no ‘unconditional promise’ when the parties can, with a wave of the pen, waive their prior handiwork. There is no ‘sum certain’ when the sum identified can be decreased or increased through purely formal (and easily accomplished) means. There is no ‘fixed maturity date,’ certainly not in the ‘reasonably foreseeable future.’59

58 See, for example, IRS Private Letter Ruling 8538034 (21 June 1985), involving the question of whether short-term discount paper issued by a US corporation to its foreign affiliates would qualify for exemption from tax at source under IRC ss881(a) and 871(g): ‘Taxpayer has represented that [its] … foreign subsidiaries will be free to exercise their own independent judgment … and will be under no obligation either to purchase or reinvest in commercial paper issued by the Taxpayer.’ The private ruling was withdrawn by IRS Private Letter Ruling 8612023 (18 December 1985), as ‘too dependent on subsequent facts’ — namely, whether the subsidiaries would ‘exercise their own independent judgment.’

59 See Laidlaw, above n39 at 2616: ‘If a transaction is controlled by related entities, the form and labels used may not signify much because the parties can mold the transaction to their will.’
4. Conclusion

Related party debt is a common, much used, element in tax shelters, cross-border tax planning (both ways), earnings stripping, reduction of the income of controlled foreign corporations through hybrid branches, and many, many other troublesome aspects of US income tax law. If they were analysed dispassionately in light of the principles that have developed for distinguishing debt from equity, ostensible loans from one entity to another would probably be viewed as equity transfers. If related party debt were not accepted as debt, or at least were not capable of giving rise to deductible interest for the purported debtor, a critical elective element in the law would be eliminated. An important set of hypothetical ‘as if’ determinations would no longer be necessary. Many sections of the Internal Revenue Code would diminish in importance, perhaps to the point of irrelevance. A number of difficult ‘look through’ rules could be simplified. Remaining tax policy choices would be more transparent to legislators, other policymakers, tax professionals, perhaps even taxpayers. The result would be a clearer reflection of income as accretion of economic power, and a more equitable income tax system.
Refugee Law: The Shifting Balance
JUSTICE RONALD SACKVILLE*

Abstract

Refugee law has only recently been recognised as a discrete legal discipline in Australia. Yet within two decades of the first statutory acknowledgement of the Refugees Convention, refugee law has become one of the High Court’s most substantial sources of work, as well as the single largest component of the work of the Federal Court.

A number of recent decisions, such as the Tampa litigation, have attracted considerable public attention. They have also tended to contribute to increasing tensions between the courts and the executive.

The High Court’s decision in Plaintiff S157/2002 v Commonwealth reaffirms the central role played by Chapter III of the Constitution and s75(v) in particular, in maintaining the rule of law. The decision effectively constitutionally entrenches judicial review of migration decisions.

But it is unlikely that Parliament has spoken its last word on the subject. It is open to Parliament, for example, to legislate in a manner inconsistent with Australia’s obligations under the Refugees Convention. That this is a realistic possibility reflects the radical changes in global social and economic conditions that have taken place since the adoption of the 1967 Protocol to the Convention. The mass movement of people seeking a better life has aroused antagonism rather than sympathy in many first world countries. While Parliament may not be able to prevent courts exercising the power of judicial review, it could well choose to limit the extent to which the Convention is part of Australian domestic law.

* Judge, Federal Court of Australia. This is a revised version of a paper delivered to the Judicial Conference of Australia’s Colloquium held in Darwin in May 2003.
1. Introduction

The emergence of refugee law as a new discipline has been one of the most striking recent developments in Australian law. This development has coincided with immigration policy attaining, not for the first time in Australian history, high prominence as a contentious political issue. The courts have accordingly become the forum for challenges to government decision-making and to legislation in a particularly sensitive area.

Since refugee law incorporates principles of judicial review of administrative action, constitutional law and international law, the stage has been set for conflict between the courts and elected governments. For the most part, this has been played out by a process of judicial interpretation of legislation (including interpretation of the Convention relating to the Status of Refugees (hereinafter Refugees Convention) insofar as it has been incorporated into domestic law), followed by a legislative response if Parliament takes the view that the courts have gone too far. Recently, however, the High Court and the Federal Court have invoked constitutional norms to read down legislation and to impose significant constraints on the power of Parliament to regulate the scope of judicial review of administrative decision-making.

On the face of it, the balance of power in this apparent contest has swung from Parliament to the courts. Yet appearances may be misleading. Ultimately it is for Parliament, regardless of Australia’s international obligations, to determine the extent to which the Refugees Convention is to form part of Australian domestic law. Indeed, there is a body of opinion that legislation has already departed from Australia’s international obligations in important respects. The recent introduction of temporary protection visas, in place of permanent visas, for successful asylum seekers reflects a growing reluctance by successive governments to accord a generous interpretation to the Refugees Convention.

One consequence of the contest between the judicial branch, on the one hand, and the executive and legislative branch, on the other, may well be a reassessment of Australia’s commitment to the Refugees Convention. If this is to occur, it is unlikely that Australia would be alone, since many first-world countries share the concern about the continued viability of international regimes for the protection of asylum seekers.

2. A New Discipline

Not so long ago, the notion that refugee law could be regarded as a discrete legal subject would have seemed very strange to an Australian lawyer. It is true that Australia has been a party to the 1951 Convention relating to the Status of Refugees since it came into force on 22 April 1954.1 But as Mary Crock has pointed out, until 1989, when Commonwealth legislation for the first time set out detailed criteria governing the grant of entry permits:
the admission or expulsion of non-citizens [including those claiming to be refugees] was regarded as a matter of ministerial prerogative and an inappropriate subject for judicial review.2

Indeed, it was not until 1980 that any Commonwealth statute made any reference to the Refugees Convention and, even then, it was for the purpose of limiting the circumstances in which the Minister could exercise a discretion to grant an entry permit to a non-citizen after his or her entry into Australia.3

A little over two decades after the first statutory acknowledgement of the Refugees Convention in Australia, a foreign devotee of the High Court’s website might gain the impression that migration law in general, and refugee law in particular, has become one of the Court’s most important sources of work. Between 1999 and the end of 2003, the full High Court heard and determined at least 27 cases concerned with migration law, most of which involved persons claiming to satisfy the Refugees Convention definition of ‘refugee’ and therefore to be entitled to protection visas.4 During the same period, migration cases, the bulk of which have involved claimants for protection visas seeking judicial review of adverse decisions, constituted over one third of the judicial caseload of the Federal Court.5 It is not surprising that this plethora of litigation has given birth to a new legal discipline.6

One reason why refugee law has developed so rapidly is that it represents the intersection of several areas of fundamental importance to the legal system, notably international law, constitutional law and administrative law. Since 1989, the Refugees Convention, a foundation stone of the post-war international legal order, has in substance been incorporated into Australian domestic law, although, the precise extent to which it has been incorporated has varied according to the constantly changing structure of the Migration Act.7 In consequence, the

---

1 In 1973, Australia also adopted the 1967 Protocol Relating to the Status of Refugees: see Ros Germov & Francesco Motta, Refugee Law in Australia (2003) 16–19 at 839. The definition of ‘refugee’ in Art 1A(2) of the Refugees Convention, as modified by the 1967 Protocol, is a person who ‘owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country; or who, not having a nationality and being outside the country of his former habitual residence is unable or, owing to such fear, is unwilling to return to it.’


4 The incidence of such cases has not diminished. During 2003 the full High Court decided ten cases involving immigration or aliens issues, of which eight were refugee matters.

5 The Federal Magistrates Court now also has jurisdiction in migration matters, pursuant to s483A of the Migration Act 1958 (Cth).

6 See, for example, Germov & Motta, n1 above; Mary Crock, Immigration and Refugee Law in Australia (1998).

7 See now Migration Act 1958 (Cth) ss36 and 65; compare Part 2, div 3, subdiv AL; compare text at nn40–43, below.
Australian courts, like their counterparts in other countries which are parties to the *Refugees Convention*, have had to construe its imprecise language. Not surprisingly, given the infinitely varied circumstances in which the *Refugees Convention* falls to be considered, it has been applied to what many critics see as an ever widening range of cases. For example, the *Refugees Convention*’s concept of a ‘particular social group’ has recently been held by the High Court to include so-called ‘black children’ in China (that is, children born outside the constraints of China’s one-child policy), women in Pakistan and homosexuals in Bangladesh. All of these decisions have the potential to increase substantially the classes of persons eligible for protection visas.

Challenges in the Australian courts by unsuccessful applicants for protection visas have provided the occasion for the elaboration and development of familiar administrative law doctrines. Thus the High Court has interpreted the requirement to comply with the principles of procedural fairness to impose what some would regard as onerous requirements on the Refugee Review Tribunal and similar bodies. While the High Court has been less tolerant of claims by disappointed applicants that the conduct of a decision-maker justifies a reasonable apprehension of bias, it is not always easy for tribunals with a heavy caseload to comply with the rigorous procedural standards prescribed by the courts in the less hectic atmosphere of an application for judicial review. To critics of judicial review of administrative action, these requirements open the way for excessive intervention by the courts into the administrative decision-making process.

Parliament has responded to the perceived generosity of the courts by enacting legislation designed to curtail the opportunities for and the scope of judicial review of migration decisions, thereby raising important constitutional questions. For example, Part 8 of the *Migration Act 1958* (Cth), enacted in 1994, deprived the Federal Court of jurisdiction to grant relief on certain grounds that otherwise would constitute jurisdictional error on the part of the decision-maker. The legislative scheme was upheld by a narrow majority of the High Court on the ground that Parliament has power, pursuant to s77(i) of the Constitution, to vest

---

8 For the approach to construction of the *Refugees Convention*, see generally *Applicant A v Minister for Immigration and Ethnic Affairs* (1997) 190 CLR 225.
9 *Chen Shi Hai v Minister for Immigration and Multicultural Affairs* (2000) 201 CLR 293.
14 Part 8 was introduced by the *Migration Reform Act 1992* (Cth) and took effect on 1 September 1994.
15 *Abebe v Commonwealth* (1999) 197 CLR 510. The result, until the repeal of Part 8 in 2001 (by the *Migration Legislation Amendment (Judicial) Review Act 2001* (Cth)), was a ‘bifurcated’ jurisdiction in migration matters, divided between the High Court and Federal Court.
jurisdiction in a federal court over part only of a controversy. More recently, Parliament’s attempt to confine judicial review of migration decisions by means of a privative clause survived a constitutional challenge, but at the price of a very narrow reading of the provision.16

3. Political Sensitivity of Refugee Law

These developments would be reason enough for refugee law to be of interest to public lawyers and to those with a particular interest in utilising the legal system to protect the interests of a vulnerable group seeking refuge in this country. But in recent years, Australian refugee law has attained greater public prominence and indeed notoriety than virtually any other area of law, except perhaps outside the criminal law. In part, this has been the product of high profile challenges to government policy, notably the Tampa litigation,17 decided in the lead up to the 2001 federal election.

In that case, the trial judge, North J, made orders directing the Commonwealth to bring ashore and release 433 asylum seekers who had been rescued from a sinking fishing boat travelling from Indonesia to Australia by the Norwegian vessel MV Tampa. The rescue had taken place about 140 kilometres north of Australia’s Christmas Island territory. The Full Federal Court, in proceedings which attracted the attention usually reserved for sensational criminal trials, in effect upheld what became known as the ‘Pacific Strategy’ to unauthorised arrivals by boat. The Court concluded that the Commonwealth, in refusing the rescuees permission to land in Australia, had acted within the executive power conferred by s61 of the Constitution.18

The Tampa litigation is however, only one illustration, albeit a dramatic one, of the peculiar political sensitivity of refugee law.19 Judicial review of administrative action always has the potential to create conflict between the courts and the executive, regardless of the political complexion of the government of the day. As McHugh J has pointed out, tensions inevitably are created by the exercise of the power of judicial review since the courts often appear to undermine executive

---

18 Ruddock v Vadarlis (2001) 110 FCR 491. Special leave to appeal was refused by the High Court.
power. The potential for tension has increased in recent times because of the expanded scope of judicial review, exemplified by the apparently ever-increasing requirements of procedural fairness and the extension of judicial review to exercises of prerogative power previously thought to be exempt from judicial scrutiny.

Although administrative law always carries with it the risk of conflict between the courts and the government of the day, there is no area that has generated more conflict than judicial review of migration decisions, especially refugee cases. In a recent paper, I identified a number of factors that have contributed to the tension between the judicial and executive arms of government. These include:

- the relative novelty of the concept of judicial review of refugee decisions which, within a short time, has converted a largely unreviewable administrative discretion into a decision-making process subjected to close scrutiny by the courts;
- the historical fact that immigration has been an especially sensitive area of public policy in Australia since and even well before federation;
- the operation of the Refugees Convention itself which, despite much talk of illegal arrivals and queue jumpers, imposes protection obligations on Contracting States towards people arriving in their territory by whatever means, provided that they can satisfy the definition of ‘refugee’ in Article 1A(2); and
- the reliance by Parliament on repeated legislative amendments to overturn unwelcome judicial decisions or to curtail the scope of judicial review, without proponents of the legislation appreciating the profound difference between their subjective intentions and the intention to be attributed to Parliament by the courts when applying well established techniques of statutory interpretation.

4. The Constitutionalisation of Refugee Law

In the same paper, and with s75(v) of the Constitution in mind, I suggested that the fate of the institution of judicial review of migration decisions was likely to rest with the High Court, rather than with Parliament. That prediction has come to

21 Id at 571.
23 One of the first enactments of the Commonwealth Parliament was the Immigration Restriction Act 1901 (Cth) which subjected potential immigrants to the notorious dictation test.
24 A point brought home starkly by Plaintiff S157/2002, above n16 at 503 (Gaudron, McHugh, Gummow, Kirby & Hayne JJ).
25 Section 75(v) provides that the High Court shall have original jurisdiction in all matters in which a writ of mandamus or prohibition, or an injunction, is sought against an officer of the Commonwealth.
In Plaintiff S157/2002 v Commonwealth, a challenge was made to s474(1) of the Migration Act, a privative clause, which on its face attempts to shield decisions of the Refugee Review Tribunal (and other decision-makers) from judicial review except on very narrow grounds. The High Court rejected the challenge to the validity of s474(1), holding that the provision, on its proper construction, does not oust the entrenched jurisdiction of the Court, conferred by s75(v) of the Constitution, to grant writs of mandamus and prohibition and injunctive relief. However, in order to avoid possible infringement of Chapter III of the Constitution, the Court gave s474(1) a very narrow construction, such that it provides no protection against review for jurisdictional error by the Tribunal. Parliament’s attempt to curtail the scope of judicial review of migration decisions therefore failed.

The major significance of the decision in Plaintiff S157/2002 v Commonwealth flows from the Court’s invocation of the Constitution as a reason for giving s474(1) of the Migration Act a narrow construction. The joint judgment implies that if the privative clause had purported to immunise decisions of the Refugee Review Tribunal against judicial review for jurisdictional error, it would fall foul of s75(v) of the Constitution. Their Honours also suggest that had a broader construction of the privative clause been adopted, the provision:

would confer authority on a non-judicial decision-maker of the Commonwealth to determine conclusively the limits of its own jurisdiction and, thus, at least in some cases, infringe the mandate implicit in the text of Ch III of the Constitution that the judicial power of the Commonwealth be exercised only by the courts named and referred to in s71.

Plaintiff S157/2002 v Commonwealth reaffirms the central role played by Chapter III of the Constitution, and s75(v) in particular, in maintaining the rule of law in Australia. In that sense, the importance of the case far transcends the High Court’s construction of the particular privative clause inserted by Parliament into the Migration Act. But the case also marks the constitutionalisation of refugee

27 Migration Act 1958 (Cth).
28 The privative clause was drafted on the assumption that it would operate subject only to the so-called Hickman proviso, whereby an administrative decision can be quashed notwithstanding a privative clause, if that decision is not a bona fide attempt to exercise the power in question, does not relate to the subject matter of the legislation or is not reasonably capable of reference to the power: R v Hickman; Ex parte Fox and Clinton (1945) 70 CLR 598 at 616 (Dixon J).
31 Migration Act 1958 (Cth).
32 Id at 47.
33 Id at 45.
35 Migration Act 1958 (Cth).
law in Australia. Instead of the tension between governments and the courts manifesting itself in differing interpretations of legislation or of the scope of executive power, which Parliament is always free to amend or clarify, the High Court has marked out a protected field of judicial review into which it appears that Parliament may not intrude.

The constitutionalisation of refugee law is not confined to the operation of s75(v) of the Constitution. In *Minister for Immigration and Multicultural and Indigenous Affairs v Al Masri*,36 the respondent to the appeal was an ‘unlawful non-citizen’ whose application for a protection visa had been rejected. He had asked to be returned to the Gaza Strip, his place of origin, but the necessary permits from transit countries could not be obtained. He therefore faced continuing detention during a period when there was no real likelihood of him being removed from Australia. The issue was whether s196(1) of the *Migration Act*37 required or authorised his continuing detention in those circumstances. Section 196(1) provides that an unlawful non-citizen detained under the ‘arrest’ provisions of s189 (as the respondent was) be kept in immigration detention until (relevantly) he or she is removed from Australia.

The Full Federal Court accepted that Parliament has the power to legislate for the detention of aliens for the purpose of their expulsion from Australia. It also accepted that legislation can empower the executive to detain an alien in custody for that purpose without infringing Chapter III of the Constitution, since such detention is neither punitive in nature nor part of the judicial power of the Commonwealth. But the Court considered that unless s196 were subject to an ‘implied temporal limitation’, a serious question of invalidity would arise. This was so because the section would then purport to authorise indefinite detention of an alien in circumstances where there is no real likelihood of his or her removal from Australia. The Court ultimately decided that the legislation permitted the respondent to be released by applying a ‘well-established principle of statutory construction concerning fundamental rights and freedoms’.38 But the reference to possible invalidity indicates that there may be significant limits to Parliament’s legislative authority on issues that governments are likely to regard as of high policy significance.

One consequence of the constitutionalisation of refugee law, particularly the central role accorded to s75(v) of the Constitution, is that the arena of conflict between governments and the courts is likely to shift. Hitherto that conflict has tended to embroil the Federal Court, as the Minister and others have argued that


37 *Migration Act* 1958 (Cth).

the Court has strayed into merits review and failed to give effect to the will of Parliament.\(^\text{39}\) Whether these criticisms have any validity is not presently important. The point is that it is the High Court, not the Federal Court, that has now substantially altered the balance between judicial power, on the one hand, and legislative and executive power, on the other, so far as decision-making in migration matters is concerned. To the extent that opprobrium is directed at courts by governments or political figures dissatisfied with what they see as judicial interference with migration policy, the High Court is now more likely to be seen as the source of the ‘problem’.

5. **International Norms and Domestic Policy**

The constitutionalisation of refugee law does not ensure, however, that the *Refugees Convention* will continue to be applied as part of Australian domestic law. Nor does it ensure that the courts will continue to be the authoritative interpreters of its provisions. The fact that Australia is a party to a treaty does not of itself incorporate the treaty into Australian law. Legislative implementation is required.

It follows that Parliament can legislate in a manner inconsistent with Australia’s obligations under the *Refugees Convention* and, in the view of some commentators, it has already done so.\(^\text{40}\) The measures that fall into this category include:

- the excision of Christmas Island, Ashmore Reef and other offshore places from Australia’s ‘migration zone’, thereby preventing persons arriving at these places from applying for visas and rendering them liable to be removed to a ‘declared country’;\(^\text{41}\)
- a statutory direction that Article 1A(2) of the *Refugees Convention* does not apply in relation to persecution for one or more of the five *Refugees Convention* reasons unless the reason is ‘the essential and significant reason’ for the persecution and the persecution involves both serious harm to the person and systematic and discriminatory conduct;\(^\text{42}\) and


\(^{41}\) *Migration Act 1958* (Cth) s5, 46A and 198A.

\(^{42}\) *Migration Act 1958* (Cth) s91R. Germov & Motta, above n1 at 190, argue that s91R, in so far as it redefines the ‘causation’ requirement, ‘is not in accordance with the proper construction or objective of the *Refugees Convention*’. 
• a further direction that in determining whether a person has a well-founded fear of persecution by reason of membership of a particular social group, the decision-maker is to disregard any fear of persecution that any other member of the family has experienced for a non-Convention reason.43

These measures signal to a more fundamental issue that is likely to play an increasingly prominent part in debates on refugee policy. That issue is whether the Refugees Convention is properly to be regarded as a product of its time, ill-suited to a world in which the mass movement of peoples fleeing persecution or simply seeking a better life is commonplace. If so, the question arises as to whether the Refugees Convention can survive in its present form either as an integral part of the international order or as a part of Australian domestic law.

The point has been raised in a recent research paper prepared for Commonwealth Parliamentarians.44 The author argues that the Refugees Convention is the product of the European experience of Nazi war-time persecutions and of the Cold War environment. She points out that most asylum seekers are now from the poorer countries of the Middle East, Asia, Africa and Eastern Europe. They are less welcome in western countries than asylum seekers from Western Europe once were.45 Moreover, the world refugee and internally displaced population has increased dramatically. Yet the core ‘non-refoulement’ obligation under the Refugees Convention46 takes no account of the impact of refugee movements in receiving countries and no provision is made for burden sharing among Contracting States.47 Likewise, the Refugees Convention gives priority to asylum seekers on the basis of their mobility and capacity to pay so-called people smugglers, while those with perhaps the greatest need remain in refugee camps. Further, so the author argues, the vague language of the Refugees Convention has been interpreted differently in different countries, with the consequence that the rates of acceptance of asylum seekers vary considerably among Contracting States.

Raw numbers give some insight into why these views have gained currency in Australia and elsewhere. According to the United Nations High Commission for Refugees (‘UNHCR’), there were 20,556,700 ‘persons of concern’ to it as at 1 January 2003. Of these, 10,389,700 were classified as ‘refugees’ and 1,014,400 as

43 Migration Act 1958 (Cth) s91S. The objective is to overturn Federal Court decisions holding that a member of a family who is at risk of persecution by reason of his or her association with another family member may have a well-founded fear of persecution by reason of membership of a particular social group (that is, the family): see, for example, Minister for Immigration and Multicultural Affairs v Sarrazola (No 2) (2001) 107 FCR 184.
45 Id at 8.
46 Article 33.1 of the Refugees Convention provides that no Contracting State shall expel or return (‘refouler’) a refugee in any manner whatsoever to the frontiers of territories where the refugee’s life or freedom would be threatened on account of his or her race, religion, nationality, membership of a particular social group or political opinion.
47 Millbank, above n44 at 12.
asylum seekers. The main countries of origin for refugee populations were Afghanistan, Burundi, Sudan and Angola. Contrary to popular belief in western countries, overwhelmingly refugee populations have found asylum in other developing countries such as Pakistan, Iran, Tanzania and Kenya. Nonetheless, some industrialised countries receive large numbers of asylum applications. These include the United Kingdom (110,700), Germany (71,100) and the United States (81,800). Moreover, the cost of processing claims and caring for asylum seekers is very considerable. It has been suggested, for example, that in 2000 the United Kingdom spent more on asylum seekers ($US2.2 billion) than the entire UNHCR budget ($US1.7 billion).

6. A New International Order?

Critics of the Refugees Convention are not confined to the ranks of politicians or administrators. Professor James Hathaway, an eminent scholar of international refugee law, argues that

the present breakdown in the authority of international refugee law is attributable to its failure explicitly to accommodate the reasonable preoccupations of governments in the countries to which refugees flee …. Apart from the right to exclude serious criminals and persons who pose a security risk, the duty to avoid the return of any and all refugees who arrive at a state’s frontier takes no account of the potential impact of refugee flows on the receiving state.

Professor Hathaway points out that much of the debate during the drafting of the 1951 Refugees Convention was devoted to considering how to protect the national self-interest of receiving states. States were not required to grant permanent residence to refugees, but merely to avoid returning them to an ongoing risk of persecution. In that sense, Professor Hathaway suggests, ‘refugee law is clearly based on a theory of temporary protection’.

Professor Hathaway makes other important observations. The 1951 Refugees Convention was formulated at a time when refugees were predominantly of European stock whose cultural assimilation was seen to be relatively straightforward. It must be remembered that the Refugees Convention in its original form was limited to persons who satisfied the definition of refugees ‘as a result of events occurring before 1 January 1951’ and contained an optional geographic limitation restricting its operation to events in Europe. It was not until

50 Millbank, above n 44 at ii–iii.
the 1967 Protocol that these restrictions were lifted. The late 1960s and the early 1970s, however, was a time of labour shortages in the developed world, particularly Europe. At that time there was, as Professor Hathaway says, a ‘pervasive interest-convergence between refugees and the governments of industrialised states’.

There has been a radical change in global social and economic conditions since the 1967 Protocol came into force. There is no longer a convergence of interest between asylum seekers and governments of advanced economies. The mass movement of people seeking a better life has aroused antagonism rather than sympathy, an attitude doubtless encouraged by the increased threat posed by international terrorism. These changes have prompted developed countries, Australia included, to adopt ‘non entrée mechanisms’ such as border controls, visa requirements for nationals of refugee-producing states, burden-shifting arrangements and forcible interdiction of asylum seekers in international waters.

Professor Hathaway argues for mechanisms to ameliorate the plight of receiving states. These, he says, should revolve around the principle that the protection obligation continues only until the refugee can return to his or her country of nationality in safety and dignity. Such an approach implies that state responsibilities may vary according to the circumstances of the receiving countries, with a greater emphasis being placed on the international community’s collective responsibility for affording protection to genuine refugees.

Australia has already proceeded along the path suggested by Professor Hathaway. Prior to 1999, all successful applicants for a protection visa became entitled to permanent residence and to the settlement support arrangements provided to refugees taken under off-shore arrangements. By regulations introduced in October 1999, provision was made for temporary protection visas for unauthorised arrivals found to be refugees. The holders of such visas received more limited benefits than those accorded permanent protection, but were eligible to apply for a permanent protection visa after 30 months provided that they were assessed at that time as still in need of protection. Substantial numbers of

53 The 1967 Protocol is not strictly an amendment to the 1951 Refugees Convention, but a separate instrument: see Minister for Immigration and Multicultural Affairs v Savvin (2000) 98 FCR 168 at 195 (Katz J). The 1967 Protocol preserved the geographical restriction for State parties to the 1951 Convention, but provided for removal of the restriction if the party so determined.
54 Migration Act 1958 (Cth) ss29(2) and 30(2); Migration Regulations 1994 (Cth) sched 2, sub-class 785.
55 The Migration Amendment (Excision from Migration Zone) (Consequential Provisions) Act 2001 (Cth) provided that holders of temporary protection visas who, since leaving their home country, have resided for at least seven days in a country where they could have sought and obtained effective protection, are not able to access a permanent protection visa in the future, but may only have access to further three year temporary protection visas while their need for protection remains: s3 and sched 2 of the amending Act. See now Migration Regulations 1994 (Cth) sched 2, item 866.215.
temporary protection visas have been granted to refugees who have arrived in Australia without authority.56

More recently, amendments to the Migration Regulations 1994 (Cth) have extended the temporary protection visa arrangements to all onshore applicants for protection, including those who enter Australia lawfully and on genuine documentation.57 The eligibility criteria for a protection visa now includes a requirement that the applicant has already been granted a temporary protection visa or a temporary safe haven visa.58 In keeping with a trend to vest greater discretionary powers in the Minister, the Minister is empowered to waive the requirement for a temporary protection visa (or the equivalent) if satisfied that it is in the public interest to do so and that the applicant entered Australia lawfully.59 According to the Minister, the changes:

will align Australia’s international protection objectives which emphasise that for most refugees the appropriate response is to provide interim protection until they can return home in safety…60

7. Conclusion

Refugee law in Australia, as in most industrialised countries, has developed extremely rapidly, over a short period. In part, this reflects world-wide trends from which Australia is not immune, despite the apparent success of the ‘Pacific Strategy’ and other measures in discouraging the flow of boat people from south-east Asia to Australia’s northern offshore territories. It also reflects the fact that the courts, including the High Court, have to grapple with a range of difficult issues, many of which are of considerable political moment. The resolution of those issues has exacerbated the underlying tensions between governments and the courts associated with judicial review of administrative action.

The constitutionalisation of refugee law, exemplified principally by S157 v Commonwealth, marks a shift in the balance between judicial and legislative powers. The High Court has identified significant limits on the extent to which Parliament can curtail the process of judicial review entrenched by s75(v) of the Constitution. Nevertheless, the ultimate authority over refugee law rests with Parliament. This has been brought home by domestic legislation that appears to

56 In 2000–2001, 4456 temporary protection visas were granted, while a further 3082 were granted in the program year to 31 May 2002: Department of Immigration and Multicultural and Indigenous Affairs, Fact Sheet 64: Temporary Protection Visas (Canberra: AGPS, 2002). For reviews of the impact of the temporary protection visa regime on visa holders, service providers and community based organisations, see Renae Mann, Temporary Protection Visa Holders in Queensland (2001); Fethi Mansouri & Melek Bagdas, Politics of Social Exclusion: Refugees on Temporary Protection Visa in Victoria (2002).
57 Migration Amendment Regulations 2003 (No 6) s3 and sched 1. See, in particular, the substituted cl866.212 in sched 2 of the Migration Regulations 1994 (Cth).
58 Id at cl866.212(1), (2) and (4).
59 Id at cl866.212(5).
depart from Australia’s obligations under the *Refugees Convention* and by the introduction of temporary protection visas as the norm for successful asylum seekers.

Australia could choose to give effect to a more restrictive refugee policy by rejecting its international obligations altogether, either by denouncing the *Refugees Convention* and *Protocol* or by enacting legislation inconsistent with the non-refoulement obligations imposed by those instruments. Unilateralism to this extent is perhaps unlikely. Nevertheless, reconsideration of the scope of the *Refugees Convention* by the international community may well be the outcome of a generally more hostile environment within first world countries to asylum seekers.

The role of the courts is to interpret and apply the *Refugees Convention*, in so far as it has been incorporated into Australian domestic law, regardless of political controversies that any given decision may generate. But those who value the *Refugees Convention* as a key element of the international order should not underestimate the significance of the changes in global social and economic conditions since the 1967 *Protocol* expanded the reach of the *Refugees Convention*. The price for continued adherence to the substance of the *Refugees Convention* may be a legislatively imposed narrowing of its scope.
Judging Refugees: The Clash of Power and Institutions in the Development of Australian Refugee Law

MARY CROCK*

Abstract

This article explores one aspect of the controversies that have surrounded asylum seekers and refugees in Australia: the conflict between the executive government and the courts over who should have the final say in status determinations and protection issues generally. This author argues that a combination of history, culture and geography has resulted in an extraordinary intimacy of political involvement in the business of immigration control — setting the groundwork for remarkable clashes with the judiciary. The article sketches the development of Australia’s jurisprudence on refugees, exploring the impact that public controversy and direct political pressure might have had on the formation of the law. The author notes that Australia’s refugee jurisprudence is recent; it is generally conservative, textual and domestic in its focus. At the same time, the author argues that the jurisprudence represents a good example of ‘globalisation’ in public international law as Australian courts have both come to consider the refugee jurisprudence of other countries and have themselves contributed to international jurisprudence on refugees.

* BA (Hons) LLB (Hons), PhD (Melb), Senior Lecturer, Faculty of Law, University of Sydney, Barrister and Solicitor (Vic and NSW). This paper reflects my years of pre-academic experience as an advocate as much as it does doctrinal research. Thanks are due to Louise Pounder for her expert editing and research assistance. I am grateful also to Ron McCallum and John Basten for their helpful suggestions and comments. The opinions expressed and any errors that remain are my own.
1. **Introduction**

Australia has had a long, but very mixed experience, with refugees. In one sense — with the exception of Australia’s indigenous peoples — it is a country made of and by refugees. If most of the earliest white settlers came here in chains, many of those who followed were people either in flight from a hostile earlier existence or in search of a better future. This is not only true of the 650,000 ‘sponsored’ refugees resettled in Australia since the end of World War II.¹ When refugees come to Australia through managed programs, Australians still prove to be very generous: witness the response given to the 4000 fugitives from Kosovo given temporary safe haven in the country in 1999. On the other hand, Australia has seen very few ‘real’ refugees in the sense of persons in need of immediate protection arriving on our door step uninvited and unvisaed: what the law refers to first as asylum seekers and (after processing) as Convention refugees.² These asylum seekers and ‘onshore’ refugees have evoked a very different response, in particular when their mode of arrival is by boat.

Australia is party to the Convention relating to the Status of refugees and its attendant Protocol,³ and as such is obliged not to ‘refoule’ or return refugees to a place where they would face persecution on one of the five Convention grounds.⁴ It has also undertaken not to punish refugees who enter the country illegally.⁵ The dilemma of on-shore refugee determinations is in the difficulties inherent in reconciling these refugee ‘rights’ with the sovereign power of the Australian government to control immigration into the country. Australia complies with its international legal obligations through its mechanisms for determining the

---


² In the 12 year period of 1989 to 2001, some 252,940 non-citizens arrived in Australia by boat or by plane without permission. This represents a little fewer than 6 such arrivals per day.

³ ‘The Refugee Convention’. The *Refugee Convention* was done at Geneva on 28 July 1951. (See Aust TS 1954 No 5, 189 UNTS No 2545, 137). The Protocol was signed on 31 January 1967, and ratified on 13 December 1973. (See, Aust TS 1973 No 37, 606 UNTS No 8791, 267). The Convention covers events causing a refugee problem before 1 January 1951, while the Protocol extends the definition to events occurring after that date.

⁴ See the *Refugee Convention*, Arts 1(2) and 33; and the Protocol, Art 1(A)(2). The *Refugee Convention* and Protocol combine to define a refugee as any person who: ‘… owing to a well-founded fear of being persecuted for reasons of race, religion, nationality or membership of a particular social group or political opinion, is outside the country of his nationality and is unable or, owing to such fear, is unwilling to avail himself of the protection of that country.’

⁵ See *Refugee Convention*, Art 31. See Guy S Goodwin-Gill ‘Article 31 of the 1951 Convention relating to the Status of Refugees: Non-penalization, Detention and Protection’, paper prepared at the request of the Department of International Protection for the UNHCR Global Consultations: <www.unhcr.ch>. (A comprehensive paper on the extent to which the Convention relating to the Status of Refugees does (and does not) permit state parties to detain and/or otherwise to penalise refugees and asylum seekers who enter a country without authorisation). See also Summary Conclusions on Article 31 of the 1951 Convention relating to the Status of Refugees – Revised, available through the UNHCR website. These conclusions were based on discussions centred on Professor Goodwin-Gill’s paper, together with written contributions that included a paper by Michel Combarnous for the International Association of Refugee Law Judges.
‘refugee status’ of non-citizens in Australia and does this (in the main) irrespective of mode of entry. It is a process, however, that can be deeply conflicted — most particularly when the determination system produces results that are at odds with the expectations of the politicians or of the general public.

This article explores the background to the judicial dramas over refugees and asylum seekers that have dominated the legal discourse in Australia in recent years. Wrought less by the refugees themselves than by our reaction to the phenomenon of fugitives on the move, these dramas provide an insight into the way Australians think about immigration control and its relationship with concepts of refugee protection.

The paper begins with a search for some explanations as to why the issue of asylum seekers and ‘on-shore’ refugees has caused friction between the government and the courts in Australia. I will argue that a combination of history, culture and geography has resulted in an extraordinary intimacy of political involvement in immigration control that has worked to the detriment of a balanced regime for refugee protection in Australia. The notion of international law conferring rights or entitlements on non-citizens who arrive without a visa or other authority to enter has been an anathema to politicians vested with the sovereign or prerogative power to determine which non-citizens enter or remain in the country. For the courts, the political focus of the refugee status determination processes set the groundwork for inevitable clashes that were exacerbated by the vagaries of particular episodes of refugee flows into Australia.

Part three of the article explores briefly the development of an Australian jurisprudence on the status and entitlements of refugees. It will be my contention that the way the case law has developed in Australia reflects closely the political pressures that have been applied in this area. The jurisprudence is recent; it is generally fairly conservative; and it is domestic and textual in its focus, with relatively little attention paid to norms of international human rights law. Having said this, the central significance of the definition of ‘refugee’ contained in the Refugee Convention means that the Australian courts have inevitably drawn from (and fed their decisions into) the international refugee jurisprudence. Indeed, curial decisions on refugee status in Australia represent one of the most striking examples of ‘globalisation’ in public international law.

2. Refugee Politics, Sovereignty and the Impact of the ‘New’ Administrative Law

Australia’s schizophrenic attitude towards refugees is often attributed to the culture of control that has always surrounded immigration to this country.6 As one of the earliest parties to the Refugee Convention done at Geneva in 1951,7 Australia was a key player in the post-war movement to establish an international

---

7 See above n3. Australia’s accession to the Refugee Convention in 1954 brought the instrument into force.
legal regime for the protection of refugees and of human rights more generally. However, our post war experience of refugees was both controlled and highly selective. It was not until the end of the conflict in Vietnam that Australia began to experience the arrival of asylum seekers in the form of boat people traveling without papers or other authority to enter the country. Geographical position and the fact that the country occupies an entire continent — sharing no land borders — has meant that Australia is one of very few countries with the ability to achieve near perfect control of immigration. Put another way, the notion that a sovereign nation should determine who enters or remains on its territory has long been seen in Australia as an achievable goal rather than empty rhetoric.

Australia received as asylum seekers very few of the thousands of boat people who fled Vietnam after the fall of Saigon.8 The crisis, nevertheless, had a profound impact on the country. While the government volunteered to accept refugees from the conflict in numbers that would literally change the cultural face of the country, it also responded by instituting the first on-shore refugee determination system. In 1978, this represented no more than a series of extra-legislative arrangements to channel refugee claims through an advisory committee. This committee would recommend the grant or refusal of refugee status to the Minister in whom Parliament had vested a simple power to grant entry permits.9

The formalisation of refugee determination procedures in Australia occurred more than 20 years after Australia became a party to the Refugee Convention and some five years after it ratified the 1967 Protocol to the Convention.10 During that period, refugee protection appears to have been an highly discretionary affair that was untrammeled by the niceties of any jurisprudence on the definition of refugee contained in Art 1(2) of the Convention.11 The significance of this is twofold. First, it can be seen that the existence of any formal procedures for the determination of refugee status in Australia is a relatively recent affair. Second, the system established in 1978 was predicated on a politician — the immigration Minister — having ultimate control over who was or was not a refugee. It will be my argument that the central involvement of politicians in the refugee determination process in Australia goes a long way towards explaining why refugees have come to be the source of so much angst in politico-legal circles. If immigration control was recognised as an incident of state sovereignty and a prerogative of government, the politicians’ understanding was (and still is) that control should be exercised by the executive arm of government and not by the courts.

8 In the late 1970s and early 1980s, small numbers of boat people arrived in northern Australia, fleeing the aftermath of the Vietnam War, which ended in 1975. Between 27 April 1976 and April 1981 2,087 Vietnamese arrived in Australia by boat, in addition to the 43,000 ‘selected’ refugees Australia agreed to accept and resettle. See Mary Crock, Immigration and Refugee Law in Australia (1998) at 127.
9 For a description of this process, see id at 127–128.
10 See above n3.
Aside from the passions and predilections of politicians too intimately involved in the refugee determination process, there are other reasons why this area of law is so politically fraught. The central tenet of the Refugee Convention is that ‘refugees’, as defined, have certain rights. The most significant of these is the right not to be ‘refouled’ or returned to a country in which they face persecution by reason of one of the five Convention grounds. The challenge of refugee law is that the corresponding obligation on Australia not to refoule a refugee — although assumed voluntarily — arises by virtue of an international instrument. The Refugee Convention is discomforting for the government because the non-refoulement obligation can be seen as a norm imposed from ‘outside’ that conflicts with Australia’s sovereign right to determine which non-citizens enter and remain in the country. There are other areas where Australian laws are shaped by international legal obligations. However, in most instances there are obvious elements of reciprocal gain in the compliance process: international trade or maritime law may be examples in point. While individual refugees have brought enormous benefits to Australia in point of fact, the idea of relinquishing sovereignty to accommodate the asylum seeker is not one that sits easily with the Australian ethic of immigration control.

It is my view that the central problem for judges charged with reviewing refugee decisions is that the Australian politicians have generally been unwilling to ‘let go’ of status determinations so as to allow the identification of Convention refugees to be a truly independent process based no more and no less than on the rule of law. From the very start there has been a tendency in immigration Ministers to personalise to themselves curial criticisms of refugee rulings.12 This may have been due in part to the rather arcane form of the relevant migration legislation — particularly during the 1980s when the Migration Act 1958 (hereinafter Migration Act) was characterised by sweeping powers vested in the Minister to grant or refuse ‘entry permits’. The personalisation process may also owe something to the language of the administrative law codified as it was by the Administrative Decisions (Judicial Review) Act 1977 (Cth). For example, when the High Court first came to rule on the interpretation of the word ‘refugee’ in 1989, it found that the misinterpretation of the law had rendered the decision made unlawful (inter alia) on grounds of ‘reasonableness’.13 That is, a decision that was so unreasonable that it could not have been made by a reasonable person.14 With a Minister who had ‘owned’ the decision legally and — one could surmise — emotionally, in putting his name to the appeal, the court’s use of this head of review was tantamount to a ruling that the Minister was an unreasonable person.

12 See the discussion of this phenomenon in Mary Crock ‘Apart from Us or a Part of Us? Immigrants’ Rights, Public Opinion and the Rule of Law’ (1998) 10 IJRL 49.
13 See Chan Yee Kin v Minister for Immigration and Ethnic Affairs (1989) 169 CLR 379, (hereinafter Chan Yee Kin) discussed below n36.
14 See Administrative Decisions (Judicial Review) Act 1977 (Cth), s5(2)(g) (hereinafter ADJR Act).
In some respects the history of refugee law in Australia represents in microcosm the changes that occurred in administrative law with the seminal reforms of the 1970s. The formalisation of refugee determination procedures coincided with the creation of the Federal Court; the passage of the *Administrative Decisions (Judicial Review) Act 1977* (Cth); freedom of information legislation; and a modest restructuring of the *Migration Act* that actually spelt out an entitlement to permanent residence to certain non-citizens who met the Convention definition of ‘refugee’. It was only a matter of time before lawyers stepped in to displace the well-connected migration agents whose business in assisting migrants was based in political deal-doing rather than the law. It will be my contention that the inevitability of the clash between the Minister and the courts over refugees was heightened by particularities of Australia’s experience with refugees.

From the perspective of my personal experience, the judicialisation of refugee law seems to have begun in earnest in 1985. Two cases were heard in that year — one in the Federal Court, the other in the High Court of Australia. In the Federal Court an application was made under the *ADJR Act* to stay the removal of a young Iranian man who was apprehended as an unauthorised arrival in immigration clearance at Melbourne airport. The man’s brother and a solicitor had been waiting for the man to arrive on the other side of customs, with a partially completed application for refugee status. Mr Azemoudeh was placed on board a plane bound for Hong Kong. Wilcox J made legal history by ordering that the plane be returned to Australia on the grounds that the decision to remove the man was rendered unlawful by the failure to consider all matters relevant to the ruling.15 The case was one of the first instances where the legal status of refugee was seen to be relevant to the lawfulness or otherwise of a ruling affecting an asylum seeker. It was a legally courageous decision because the prevailing wisdom was that immigration applicants — whether refugees16 or otherwise — had no legal right to a hearing before being expelled or excluded from Australia. The rules of ‘natural justice’ did not apply to these people.17

The second case in 1985 was instituted to question the refusal of refugee status to a man who had fled to Australia from Irian Jaya at the height of the take-over of that country by Indonesia. Ran Rak Mayer sought to challenge the ruling by the Determination of Refugee Status (DORS) Committee, which led to a recommendation to the Minister against the grant of refugee status. The High Court ruled in that case that the processes of the DORS Committee could be reviewed along with the ultimate decision by the Minister. The Court found that although the Committee was established outside the legal framework of the

---

15 See *Azemoudeh v Minister for Immigration and Ethnic Affairs* (1985) 8 ALD 281. In the event, the Minister did not comply with the order made, although Mr Azemoudeh was taken off the flight in Hong Kong and permitted to apply for refugee status there.
17 See *Salemi v Mackellar [No 2]* (1977) 137 CLR 396.
Migration Act, its deliberations were part of a process of decision making that was made ‘under an enactment’ for the purpose of s3 of the ADJR Act.¹⁸

Mayer’s case was followed one year later by Kioa v West,¹⁹ the case that did more than any other to entrench the rules of procedural fairness as the cornerstone of administrative justice in Australia. The cumulative effect of these rulings was an explosion of migration cases in the Federal Court. In the space of five years, the Federal Court used the grounds of review in the ADJR Act to turn the notion of administrative discretion on its head. It pronounced unlawful decisions where the administrator had not taken into account matters that the court considered to be of central importance, or where the court considered that the individual involved had not been given an adequate hearing. Put simply, political disquiet with the migration jurisprudence from the 1980s explains the decision in 1989 to remove the broad discretions that had characterised the Migration Act, replacing the simple powers with the complex matrix of Act and regulations with which we are now familiar.

In 1989 non-citizens seeking residence on family or employment-related grounds gained a right to review decisions before a tribunal empowered to take oral evidence. In contrast, refugee status decisions were still being made on the papers, and behind closed doors: claimants had no right to an oral hearing on appeal. The tiny number of refugee claimants seeking protection in Australia in the 1980s may explain why the system for determining refugee status was largely left intact during this first and greatest revision of the migration legislation. However, as events were to unfold in and after 1989, there was also some evidence that the decision to leave refugees out of the administrative law reforms to the Migration Act was deliberate and politically significant.

As explored in the following section, 1989 was a watershed year in the development of an Australian jurisprudence on the definition of refugee. It was also a year that put refugees on the map of political and public consciousness in a way that had not been experienced previously. The year was dominated by stories of conflict and drama in the Asia Pacific Region. Cambodia was in turmoil with the defeat of the Khmer Rouge and the withdrawal of the Vietnamese under the auspices of a United Nations peace plan. In June, after months of euphoric reporting about the imminent demise of communism in China to the forces of the Pro-Democracy movement, the Chinese government instituted a brutal and comprehensive crackdown on its citizenry. In September the High Court delivered

---

¹⁸ In Mayer v Minister for Immigration and Ethnic Affairs (1985) 157 CLR 290, the Court recognised s6A(1)(c) of the Migration Act as the source of the Minister’s power to grant refugee status. It held (at 302) that refugee status decisions were judicially reviewable under the ADJR Act and that reasons for such decisions could be sought under s13 of that Act. The case confirmed the Minister’s obligation to consider claims for refugee status and established that the standard governing the determination of refugee status was that set down in the Refugee Convention.

its ruling in *Chan Yee Kin*, confirming that this particular fugitive from China did, indeed, meet the definition of refugee. Then, in November, the boats began to arrive — carrying fugitives first from Cambodia and later from China.

It is at this point that the political orientation of Australia’s refugee determination system became both apparent and problematic. The overwhelming impression I have of this period is of the intimacy of the politicians’ involvement in every aspect of the emergent refugee ‘problem’ facing Australia. There was little or no sense of respect or reverence for a legal regime for the protection of refugees founded in independent (a-political) status adjudication. As time wore on, any notion of national or institutional compassion for the dispossessed all but disappeared. Prime Minister Hawke shed public tears over the fate of the Chinese students in Tiananmen Square, vowing that no Chinese student in Australia would be forced to return home against their will. A short time later, however, the Cambodian fugitives were told that Bob was ‘not their uncle’, and that they should return to Cambodia because the United Nations’ peace plan had made it safe to return. Prime Minister Hawke and Foreign Minister Evans labeled the fugitives ‘economic refugees’.

The Labor government’s response to the Cambodian and Chinese boat people of the early 1990s set the course for the policies and institutional hostilities that continue to this day. This period also saw the emergence of an extensive network of refugee advocates in Australia. The hapless Cambodians were the subjects of a seemingly interminable number of applications for judicial review. They were at the heart of what I have described as the tit-for-tat legislative campaign that saw each major refugee ‘win’ in court met by ‘remedial’ legislation. One infamous example of this was the hurried introduction of the mandatory detention provisions later challenged in *Chu Kheng Lim v Minister for Immigration and Ethnic Affairs*. Another was the legislation to cap any compensation payment to an asylum seeker detained unlawfully at one dollar a day. Under the Ministries of Gerry Hand and Nick Bolkus, the asylum seekers were moved around Australia, often for no apparent purpose other than to place physical distance between the detainees and their lawyers. It was the era that saw the construction of the Port

20 *Chan Yee Kin*, above n13.
21 See *Migration Regulations* 1994 (Cth) sched 2, subcl 437, which created a special visa subclass for these people. See the discussion in Crock, above n8 at 131.
22 See, for example, Prime Minister Hawke’s interview with Jana Wendt on *A Current Affair* (6 June 1990).
24 The provisions were introduced when it became clear that O’Loughlin J in Adelaide was going to order the release of the Cambodian detainees on the ground that they were being held without lawful warrant. See Mary Crock, ‘A Legal Perspective on the Evolution of Mandatory Detention’ in Mary Crock (ed), *Protection or Punishment: The Detention of Asylum Seekers in Australia* (1993) at 34.
25 (1992) 176 CLR 1 (hereinafter *Chu Kheng Lim*).
26 See *Migration Amendment Act (No 4)* 1992 (Cth).
Hedland detention facility and the institution of ‘incommunicado’ detention for newcomers.  

If refugees were left out of the administrative reforms of the migration legislation in 1989, there is evidence to suggest that the decision to eventually grant asylum seekers access to oral appeals was forced on the government by the Courts. The old appeal system was based entirely on claimants reducing their cases to writing. In a number of instances, adverse decisions were challenged on the grounds that the failure to grant an oral hearing constituted a failure to accord natural justice or procedural fairness. In these cases, the Federal Court declined to accept that the rules of procedural fairness necessarily entailed a right to an oral hearing. Instead, the court focused on the nature of the ‘hearing’ given and on the question of whether, as a matter of fact, the claimant had been given a proper opportunity to present his or her case, and to answer this or that adverse, critical and significant matter. The courts professed not to be concerned with the modalities of the hearing process; just with the issue of whether the claimant had been ‘heard’.  

Faced with the degree of judicial scrutiny that flowed from easy access to court and the simplification of judicial review procedures, it was inevitable that practicalities would force the government to introduce oral hearings for refugee appeals.

It is my personal opinion that refugee law in Australia has never recovered from the process of brutalisation that occurred during the period of the late 1980s and early 1990s. The mentality of answering court actions with legislation has remained and intensified. The result has been an incremental hardening in laws and policies. On the one hand the changes have seen Australia drift ever further from the path of the good international citizen. In our continued insistence on a system of mandatory detention, in our abuse of the human rights of refugee children and of families, Australia has become increasingly blatant in its disregard for the norms of international human rights law. The changes have also had an effect, however, at a more immediate jurisprudential level. In 1989, the codification of migration laws saw an abrupt removal of the sweeping discretions that had characterised the legislation.  

Within the space of a few years, the power of lower level officials to respond flexibly and with humanity to individuals in situations of need was all but removed.

---

27 This practice continues today and means that boat (and plane) arrivals are taken into detention and screened by departmental officials to determine whether they ‘engage Australia’s protection obligations’. Only those deemed to have a valid claim for refugee status are permitted access to a lawyer. Individuals held during this screening process are kept apart from other asylum seekers who are having refugee claims assessed. The rationale is that separation is necessary to prevent the newcomers from being coached by longer term detainees. See Mary Crock, ‘A Sanctuary Under Review: Where To From Here For Australia’s Refugee and Humanitarian Program?’ (2000) 23 UNSWLJ 246 at 273–4.


29 See Migration Legislation Amendment Act 1989 (Cth); Migration Regulations 1989 (Cth).

30 For more commentary on this point, see Sean Cooney, The Transformation of Immigration Law (1995).
The removal of general safety net provisions has ramifications at two levels. First, it has acted like a vortex to channel power back to the Minister and to no one else — increasing the personal power of the incumbent of that office to unprecedented levels. Second, it has meant that individuals facing serious personal problems (if returned to their country of origin) have had no option but to seek recognition as refugees. In simple terms, the definition of refugee in Australia in recent years has had much more protection work to do than is the case in other countries where a general discretion to grant residence of humanitarian grounds has been retained. It is difficult to quantify the impact, if any, that these developments might have had on the judicial review of refugee rulings. For judges imbued with the vision of curial review as a protective shield for the rights of the individual, it would be interesting to study whether the narrowing of avenues for persons at risk has lead to more expansive interpretations of refugee law in this country.

Refugee claimants in Australia have had the right to appeal negative status determinations to an independent administrative tribunal since 1993. However, the change seems to have done little to diminish the government’s micro-management approach to refugee protection. The present Minister has been very vocal in his criticisms of both the Refugee Review Tribunal and the courts when rulings by either body conflict with his understanding of the law. As Shadow Immigration Minister in 1992, Mr Ruddock supported the first attempt to constrain the judicial review of migration decisions by referring specifically to the High Court’s ruling in Chan Yee Kin. He said:

When we look at the creative way in which the High Court of Australia got into the business of determining refugee claims, when it was always intended that these should be administrative matters dealt with by the government of the day, we can appreciate that the government by allowing the ADJR Act to continue in this area was creating a rod for its own back. It has always seemed to me and I have argued this strongly, that the role of the courts collectively in this area has brought about a significant problem for the government of the day.

In 1997, the Minister warned Refugee Review Tribunal members publicly that they should not expect their contracts to be renewed if they purported to ‘re-invent’ the definition of refugee (by recognising that a woman victim of domestic violence could be a refugee). (Unelected) judges interpreting the law in a manner contrary

31 See Migration Reform Act 1992 (Cth), Pt 4A.
33 See article and editorial The Canberra Times (27 December 1996) at 14. At the time of the reappointment process in 1997, the recognition rate for refugee claims in the tribunal plummeted from 17 per cent to less than 3 per cent. See evidence supplied by Mr Mark Sullivan, Deputy Secretary of the Department, to the Senate Legal and Constitutional Legislation Committee. See that Committee’s report: Consideration of Migration Legislation Amendment Bill (No 4) 1997: Minority Report (Canberra: AGPS, 1997) at 45–46.
to the (elected) Minister’s understanding have been charged with subverting the ‘will of the people’.

Attitudes such as this have inevitably placed the government at loggerheads with the courts in a country where the rule of law is founded in judicial exegesis of written text, legal precedent and principle that are not infrequently at odds with the understandings of individual politicians.

3. The Development of a Refugee Jurisprudence in Australia

The challenge for the courts is plain. On the one hand, the legitimacy of any unelected judiciary is dependent on judges ensuring that they do not stray too far from the shared understandings and values of the polity. The ‘dualist’ vision of the intersection between international obligation and domestic law has permitted the courts to accommodate political sensitivities about many aspects of international human rights law: the ruling in *Chu Kheng Lim* upholding the legal validity of the first mandatory detention regime is an example in point. On the other hand, the cornerstones of refugee law are set by the terms of the *Refugee Convention*. Incorporated into Australian law either directly or by reference, the standards are now the subject of an elaborate international and an increasingly elaborate national jurisprudence. This body of law increases the pressures on the judiciary when faced with a Minister who is unabashedly assertive about his preferred interpretation of the Convention provisions.

Although reaching back little more than one decade, the jurisprudence on refugee law in Australia has now developed to the point of some sophistication. From the perspective of an observer and sometime participant in the evolution of the law in this area, four points stand out in the developing case law. First is the pragmatic, reactive approach taken by the courts; second, the narrow textual focus of many of the decisions; third, the lack of attention paid to broad norms of international law; and, finally, the (belated) intrusion of international jurisprudence into the judicial discourse on refugees.

It is beyond the compass of this article to examine the evolution of refugee law in Australia in any detail. However, there are a few key rulings that are deserving of mention. The first case in which the High Court considered in any detail the definition of ‘refugee’ contained in Art 1(2) of the *Refugee Convention* and Protocol was *Chan Yee Kin*. This is the case that set the now famous ‘real chance’ test. The court confirmed that an individual must show both a subjective fear of persecution upon return to his or her country of origin and demonstrate from an

---

34 See The Honourable Philip Ruddock, ‘Immigration Reform: Unfinished Agenda’ (speech delivered at the National Press Club, Canberra, 18 March 1998) at 7 of the electronic transcript: <http://www.immi.gov.au/kits/the_facts/speech.htm> (19 February 2004), in which he stated ‘Only two weeks ago a decision to deport a man was overturned by the Federal Court although he had been convicted and served a gaol sentence for possessing heroin with an estimated street value of $3 million. Again, the courts have reinterpreted and rewritten Australian law—ignoring the sovereignty of parliament and the will of the Australian people. Again, this is simply not on.’ [Emphasis added.]


36 *Chan Yee Kin*, above n13.
objective viewpoint that fear is ‘well founded’ — in other words that there be a ‘real chance’ of persecution if returned. The case was decided at the height of the dramas surrounding the repression in China of the pro-democracy movement. I have argued that more than one aspect of the ruling suggests that the court was moved by the obvious plight of dissident Chinese nationals. 37 Another feature of the ruling, however, was the refusal by the Court to enter into detailed debates about the proper interpretation of the Convention definition. In the leading judgment, McHugh J dismissed suggestions of major difference between the interpretations that were preferred by Courts in England and North America. He offered the ‘real chance’ test as a distillation of the prevailing international jurisprudence on the ground that all the tests pointed to similar outcomes. 38

If the ruling in that instance was adverse to the administration, in later cases judicial pragmatism led to less expansive results for the refugee claimants. For example, in Wu Shan Liang, 39 the High Court roundly rebuked the Federal Court for scrutinising too closely the reasoning methods adopted by the Refugee Review Tribunal. The High Court rejected as heresy the notion that tribunal members who speculated about the likelihood of persecution fell into legal error because such a process implied a ‘balance of probability’ rather than a simple ‘real chance’ test. 40 The ruling reflected changes to the Migration Act that reduced the scope for judicial oversight of refugee decisions. 41 However, it also marked the beginning of a period when the public discourse on refugees was becoming more negative. Sympathy for the Chinese students was giving way to near hysteria about the continued arrival of boats from China — Wu Shan Liang was actually a class action brought on behalf of the asylum seekers from one such boat. Particular concerns were being raised about asylum seekers claiming to be fugitives from China’s ‘One Child’ policy. The assertion that victims of this policy were refugees because they feared persecution by reason of their membership of a ‘particular social group’ — parents wishing to have more than one child — raised crude spectres of refugee law becoming the conduit for millions of fugitives from the world’s most populous country flooding into Australia. 42

37 See Crock, above n12 at 54–55.
39 Minister for Immigration and Ethnic Affairs v Wu Shan Liang (1996) CLR 259.
40 See Mok Guok Bouny (1994) 55 FCR 375, where the Full Federal Court held that the real chance test did not allow the RRT to engage in a process of weighing up evidence with a view to determining the likelihood of future persecution. It found that the use of expressions such as ‘I give greater weight to’ suggested that the Tribunal was assessing refugee claims on the ‘balance of probabilities’.
41 Section 22A of the Migration Act was amended to provide that: ‘If the Minister is satisfied that a person is a refugee, the Minister may determine, in writing that the person is a refugee’. The effect was to change the determination of refugee status from a process based on an objective test to a subjective (opinion based) ruling.
42 See, for example, ‘Editorial’ Sydney Morning Herald (26 February 1997) at 14; Padraic McGuiness, ‘Illegal Immigrants Cost us Dearly in Legal Aid’ Sydney Morning Herald (27 February 1997) at 15. See also below n47.
The case law involving fugitives from China’s One Child Policy shows an interesting divergence in judicial approach. It is in these cases that the divide first becomes apparent between those judges who see the Refugee Convention as an instrument for the protection of human rights, and those who see the document as a creature of history and compromise created for the purpose of regulating and controlling refugee flows. The legislative response to the cases also demonstrates vividly which of these two approaches has been preferred by the government.

In Applicants A and B v Minister for Immigration and Ethnic Affairs, Sackville J concentrated in his first instance judgment on the human impact of the Chinese policies and on the ease with which dissentients could be identified (and penalised) as members of a group within Chinese society. That case involved a Chinese couple who fled to Australia when the wife fell pregnant without first gaining the requisite permission. Giving birth very shortly after arriving in Australia, both the wife and husband claimed that they would be forcibly sterilised if returned to their village of origin. His Honour ruled that sanctions of forced abortion and sterilisation did amount to ‘persecution’, even though the policies were of general application. He also ruled that the couple’s fears of persecution were ‘by reason of’ their membership of a particular social group. The government responded to the judgment by immediately introducing into Parliament amending legislation stipulating that the fertility control policies of foreign governments cannot be used to found a claim that a person belongs to a particular social group for the purposes of making out a claim for refugee status. Labor Senator McKiernan urged passage of the Bill with warnings that Australia may have to institute an interdiction program to forestall an invasion by would-be Chinese parents. He said:

I would anticipate that hundreds of thousands of people from China and some other Asian countries will shortly be making plans to get to Australia. They will hear, if they have not already heard, that the numbers are in the Senate to defeat the Bill …. Turning boats, that carry illegal migrants to Australia, around at sea, may be the only way to stop the flood gates opening and protect Australia in the long run.

The amending legislation was never put to a vote — in large measure because Sackville J was overruled on appeal. The Full Federal Court and, later, a majority in the High Court held that the couple did not meet the Convention definition of

45 This finding was upheld by the High Court, although the Full Federal Court held against the applicants on this point: Minister for Immigration and Ethnic Affairs v Respondents A and B (1995) 57 FCR 309.
46 Migration Legislation Amendment Bill (No 3) 1995 (Cth), cl2.
47 Senator Jim McKiernan, Interdiction May have to be Considered Media Release (2 March 1995).
refugee because ‘social group’ could not be defined solely by a shared fear of persecution — there had to be some other cognisable factor through which persons external to a group could identify the group as such. The appeal judges — with the notable exceptions of Brennan CJ and Kirby J — preferred a structured, almost mechanistic, reading of the Refugee Convention definition of refugee. Although obscurities in the textual definition were identified, these were resolved through a detailed exploration of the historical background to the 1951 Convention and that instrument’s travaux preparatoires (preparatory works). On this basis, the majority concluded that the 1951 Convention was never intended to cover all persons in situations of need, justifying its own restrictive interpretation of the instrument.49 On the other hand, both Kirby J and Brennan CJ examined the Preamble to the Convention and preferred to focus on the instrument’s overriding humanitarian purpose. Against this background, their Honours found congruence between the definition of refugee and the applicants’ situation.50

Put simply, the judges who have attempted to shape their interpretation of the Refugee Convention so as to conform to the human rights spirit of the instrument have been in the minority.51 The ‘One Child policy’ cases aside, there have been other instances in which the High Court has chosen to eschew a human rights approach so as to restrict the reach of Convention protection in Australia. One notable example is the Court’s refusal to recognise as refugees, fugitives from countries in the grip of civil war, where it is not possible to identify any form of civil government or formal ‘state’.52 These cases also demonstrate the second ‘trend’ apparent across Australia’s refugee jurisprudence: the tendency to focus on text rather than context. In Ibrahim’s case, Gummow J in the majority devotes considerable energy to divining the intent of the original drafters of the Refugee Convention, concluding that the instrument was not designed to assist victims of general unrest or anarchy.53 The majority preferred what is described as the ‘accountability’ theory of refugee law, formulated by the House of Lords in Adan v Secretary of State for the Home Department.54 Lord Lloyd of Berwick said:

[W]here a state of civil war exists, it is not enough for an asylum-seeker to show that he would be at risk if he were returned to his country. He must be able to show what [counsel for the Secretary of State for the Home Department] calls a differential impact. In other words, he must be able to show fear of persecution for Convention reasons over and above the ordinary risks of clan warfare. [Emphasis added.]

49 See, for instance, Applicant A v Minister for Immigration and Ethnic Affairs (1997) 190 CLR 225 at 324 (McHugh J) and 277–78 (Gummow J).
50 Similar divisions emerge in the case of Minister for Immigration and Multicultural Affairs v Khanar (2002) 210 CLR 1 at 17 (McHugh & Gummow JJ) and 37 (Kirby J), with McHugh and Gummow JJ again emphasising the limited scope of the Convention, and Kirby J stressing its humanitarian objects.
51 See Crock, above n23.
52 See Minister for Immigration and Multicultural Affairs v Ibrahim (2000) 204 CLR 1.
53 Id at 45–50.
54 [1999] 1 AC 293.
55 Id at 311. (Lord Goff of Chieveley, Lord Nolan, & Lord Hope of Craighead agreed.)
On this view, the Refugee Convention is conceptualised as a device for overcoming the failure by a (specified) state to protect its nationals. The opposing view of the ‘protection’ theorists is that the instrument is designed first to protect persons in need who do not have the protection of any state.56

The tendency in the Australian Courts to focus on text rather than context is reflected more broadly in a tendency to concentrate very much on relevant domestic legislation to the exclusion of norms of international law. It is neither possible nor perhaps desirable to ‘prove’ that this tendency reflects domestic political pressures brought to bear on the courts. Having said this, it is in the most politically charged refugee cases that the trend towards juridical introspection is most marked. In the early 1990s, the controversy surrounding the Cambodian boat people came to a head in litigation brought with the assistance of public interest advocates to challenge both the initial refusals to grant refugee status and the decisions to hold the asylum seekers in detention. As noted earlier, the early skirmishes in the Federal Court lead to the hasty passage of the first mandatory detention provisions in the Migration Act — which lead in turn to the constitutional challenge to the provisions in Chu Kheng Lim.57

The High Court upheld the constitutionality of the detention laws by focussing narrowly on the alienage of the detainees; the nominal temporal constraints placed on detention; and the breadth of the power in the Constitution to legislate with respect to aliens. The status of the detainees as asylum seekers rates no mention. No reference is made to the constraints imposed by the Refugee Convention on the detention of ‘refugees’. For the refugee lawyer, there is special irony in the central finding that the detainees were not being held illegally because they were free to go ‘home’ at any time of their choosing: they were ‘voluntary’ detainees. The broader norms of international human rights law were dismissed as non-binding, if not irrelevant to the determination of the litigation. Article 9(1) of the International Covenant on Civil and Political Rights prohibits the ‘arbitrary’ detention of any person. In spite of Australia’s signature and ratification of this instrument, it was of no assistance to the Cambodians because of the Australian parliament’s failure to enact the relevant provisions into the domestic laws of the country.58

The extent to which the High Court’s reading of the legality of the first mandatory detention laws was at odds with international law emerged when the Cambodians took their case to the UN Human Rights Committee. In A v Australia,


57 Chu Kheng Lim, above n25.

that Committee rejected the argument that detention sanctioned by Parliamentary process cannot be ‘arbitrary’. Preferring to focus on the circumstances surrounding detention and the possibilities for release, it rejected the fiction that asylum seekers can be ‘voluntary’ detainees. The Committee made the simple point that the detainees would plainly and obviously ‘go home’ if such a course were safely open to them. The very essence of the status of refugee is that refugees cannot ‘go home’ without risking persecution.

The detention jurisprudence from the UN Committee did not prevent the Australian courts from adhering to the same personal choice fictions when faced with the most recent — and most politically explosive — incident involving asylum seekers and refugees. In spite of an initial victory before a single judge of the Federal Court, challenges made to the government’s actions in the Tampa affair also failed because of the ‘domestic’ focus of the courts. The (public interest) applicants were acknowledged as having standing to bring an action for orders in the nature of habeas corpus so as to seek the release of the Tampa rescuees. However, the government’s refusal to allow any access to the persons taken on board the Tampa meant that the applicants were unable to gain instructions from the rescuees for the purpose of mounting a legal challenge under the Migration Act. Most significantly, the Full Federal Court held that the applicants had no standing to seek a writ of mandamus to compel the Minister to act in accordance with the law.


60 On this point, see Amuur v France (1996) 22 EHRR 533, discussed below.

61 The advocates succeeded at first instance before North J in the Federal Court, who held that the asylum seekers were being detained by the Australian authorities in circumstances where there was no basis in Australian law for the action being taken. See Victorian Council for Civil Liberties and Ors v Minister for Immigration, Multiculturalism and Indigenous Affairs (2001) 110 FCR 452 (hereinafter VCCL v MIMA).

62 The Federal Court, both at first instance and on appeal, acknowledged that the liberty of the person is a human right of such fundamental importance that any person has the right under the Common Law to challenge the legality of another’s detention. See Victorian Council for Civil Liberties and Ors v Minister for Immigration, Multiculturalism and Indigenous Affairs (2001) 110 FCR 452 at 469; and Ruddock v Vadarlis (2001) 110 FCR 49 at 509 (Black CJ) and 518 (Beaumont J). See also Waters v Commonwealth of Australia (1951) 82 CLR 188 at 190; Truth About Motorways Pty Ltd v Macquarie Infrastructure Investment Management Limited (2000) 200 CLR 591 at 600 (Gleeson CJ & McHugh JJ), 627 (Gummow J), and 652–653 (Kirby J); and Clarkson v R [1986] VR 464 at 465–466.

63 The applicants’ substantive claim was based on an argument that the Migration Act 1958 (Cth) operated to require the landing of the rescuees. Eric Vadarlis asserted that the status of the rescuees as unlawful non-citizens meant that they had to be taken into immigration detention within Australia pursuant to s189 of that Act. In the alternative, the Victorian Council for Civil Liberties argued that s245F required the government to bring the rescuees to the mainland of Australia, where they would then be entitled to lodge formal claims for refugee status pursuant to s36 of the Act. See Ruddock v Vadarlis, above n62 at 506–507. Eric Vadarlis also argued that the refusal to allow him access to the rescuees constituted a breach of his implied constitutional freedom of communication. He sought an injunction and mandamus to allow him to give legal advice to the rescuees. See VCCL v MIMA, above n61 at 489–490; and Ruddock v Vadarlis, above n62 at 530.
Even though the applicants had standing to question the detention of the Afghans taken in by the *Tampa*, it did not follow automatically that the writ of habeas corpus would run. The remedy was dependent on the applicants demonstrating that these people were being detained and that their custody was unlawful. In the Full Federal Court, the majority held against the applicants on both these counts. French J, with whom Beaumont J agreed, relied on the reasoning of the High Court in *Chu Kheng Lim*. Again — and just as counter-intuitively as in the earlier case — the court held that the *Tampa* asylum seekers were not being ‘detained’ at law because they were free to travel anywhere they wished (except to Australia). Both North J at first instance and Black CJ on appeal disagreed strongly with this characterisation of events. In a persuasive dissent, Black CJ examined the manner in which the European Court of Human Rights treated the same issue in *Amuur v France*. That case involved four Somali asylum seekers who were kept for 20 days in the transit zone of the Paris-Orly airport. The argument was made that the asylum seekers were not detained because they could have left at any time by agreeing to return to Syria, from whence they had traveled to France. The Court said:

The mere fact that it is possible for asylum seekers to leave voluntarily the country where they wish to take refuge cannot exclude a restriction on liberty…. Furthermore, this possibility becomes theoretical if no other country offering protection comparable to the protection they expect to find in the country where they are seeking asylum is inclined or prepared to take them in.

The *Tampa* judgments are redolent with political stresses operating on the presiding judges. North J at first instance was the only judge to acknowledge that the group in question most probably included Convention refugees as most were from Afghanistan, where the feared Taliban were still in power. Yet the judge chose the neutral term of ‘rescuees’ to describe *Tampa*’s hapless human cargo. Dissenting in the Full Court, Black CJ agreed with the substance of North J’s rulings. However, his carefully reasoned judgment sticks closely to legal principle, assiduously avoiding any emotive descriptions of the people behind the action. In

---

64 See *Chu Kheng Lim*, above n25. For an account of this case, see Crock, above n58.
65 *Ruddock v Vadarlis*, above n62 at 548.
66 *Ammur v France*, above n60.
67 Id at 558. See *Ruddock v Vadarlis*, above n62 at 510 (Black CJ).
68 His Honour said: ‘It is notorious that a significant proportion of asylum seekers from Afghanistan processed through asylum status systems qualify as refugees under the Convention relating to the Status of Refugees (1951) (the Refugees Convention). Once assessed as refugees, this means that they are recognised as persons fleeing from persecution in Afghanistan. While such people no doubt make decisions about their lives, those decisions should be seen against the background of the pressures generated by flight from persecution.’ See *VCCL v MIMA*, above n61 at para [67].
69 Most of the *Tampa* rescuees were Afghani nationals. Before the US intervention in Afghanistan led to the defeat of the ruling Taliban in late 2001, over 80 per cent of Afghan asylum seekers in Australia were gaining recognition as refugees. Of the 130 *Tampa* asylum seekers accepted by New Zealand, all but one were recognised at first instance as refugees and offered permanent resettlement.
contrast, the second majority judgment — by Beaumont J — is replete with a sense of urgency, heightened perhaps by the unmentioned spectre of the terrorist attacks in America that occurred on the same day that North J handed down his ruling. Beaumont J affirmed as pre-eminent the government’s power to determine which non-citizens could enter and remain in Australia, underscoring passages and words. His conclusion — that an alien has no right to enter Australia — is placed quite literally in bold print. The effect is to emphasize and re-emphasize the outsider status of the rescuees. The word ‘alien’ appears no less than 27 times in the 30 paragraphs of his judgment.

Perhaps the greatest evidence in the Tampa litigation of the political pressures at play, however, are the obiter dicta comments of French J on the question of whether the government’s actions were rendered unlawful by the fact that there was no legislative basis for what occurred. His Honour suggested that the nature of the executive power conferred on the government under the Australian Constitution may be such that legislation was not needed to render lawful any actions taken to protect Australia’s borders. Amidst the flurry of legislative change on 26 September 2001, French J’s comments did not go unnoticed by government drafters. The amendments to the Migration Act included a new s7A, which confirms the power of the executive to act outside of any legislative authority. The new section reads:

> The existence of a statutory power under this Act does not prevent the exercise of any executive power of the Commonwealth to protect Australia’s borders, including, where necessary, by ejecting persons who have crossed those borders.

In referring to ‘persons’, s7A makes no distinction between citizens and foreigners. This provision raises very interesting questions about the reach of the executive power in matters of border protection. It is beyond the scope of this article to explore this issue in any detail. Nevertheless, the concept of the executive having completely untrammelled power over either or both non-citizens and citizens who have crossed the border and entered the country, sits uneasily with the balance of powers established by the Australian Constitution.

The refugee story in Australia undoubtedly reached a very low point in 2001. However, it is another thing to see the Tampa judgments as representative of the trend in Australian refugee jurisprudence. A longer-term view of the refugee cases during the 1990s suggests that the courts in Australia are becoming increasingly familiar with both the terms of the Refugee Convention and Protocol and with comparative international jurisprudence on these instruments. In cases where political pressures have not been so pressing, the tendency has been for the Australian courts to move with the international trends, even where to do so has placed them at loggerheads with the government.

The High Court’s treatment of the ‘One Child policy’ and domestic violence cases are examples in point. Both involved situations where the Minister had made

---

70 Ruddock v Vadarlis, above n62 at 543–544 (French J). Given his Honour’s findings on the detention point, these comments probably constitute obiter dicta.
it very plain as to how he preferred the law to be interpreted. However, in both instances, the passage of time and events eventually operated to reduce or remove the element of public angst about the issues involved.

While the first cases involving fugitives from China’s One Child policy excited a great deal of controversy, it quickly became apparent that the country was not being flooded with refugee parents from China. On the other hand, the Australian public showed a degree of discomfort when presented with evidence that a failed refugee claimant returned to China in a heavily pregnant state was either forced or coerced into having an abortion upon her return. In April 2000, the High Court affirmed that children born in contravention of China’s One Child policy — known as ‘hei haizi’ or ‘black’ children — can be refugees. Outside their country of origin, faced with fines and other penalties restricting access to education and other public services, ‘black’ children were recognised by the High Court as belonging to a particular social group. The court ruled that this grouping was sufficient to render the children refugees because of the immutability of the factors predetermining the (persecutory) treatment that the children would receive. One irony of the court’s ruling in Chen Shi Hai is that the child in question came to Australia on the same boat as the ill-fated pregnant Chinese woman. In the woman’s case, not only was her pregnancy insufficient to gain her protection in Australia but she was also returned to China with an older daughter born out of wedlock in Australia — a child who was indisputably a ‘black’ child in China. If the mother did not meet the stringent test for refugee status set by the High Court, there is every indication that her little daughter was a Convention refugee.

In Khawar’s case, the High Court again adopted an expansive interpretation of the Convention definition of ‘refugee’ to recognise that a Pakistani woman fleeing domestic violence could be a refugee. On this occasion the Court held that women in Pakistan could constitute a cognisable social group. The court further held that the (systemic) failure of the Pakistani government to intervene so as to protect women in Pakistan from domestic violence was sufficient to convert the serious harm they suffered into ‘persecution’ for the purposes of the Refugee Convention.

While it is possible to read the decisions in Chen Shi Hai and Khawar in strict doctrinal terms that are consistent with the earlier rulings in Applicants A and B and Ibrahim, the outcomes in the cases are dramatically different. In both Chen Shi Hai and Khawar, the Court’s interpretation of the words of Art 1(2) was neither inevitable nor incontrovertible — a fact brought out strongly by the Minister in his

71 The woman was known as ‘Ms Z’. Her case was both the subject of an inquiry personally instituted by the Minister for Immigration and Multicultural Affairs (a report was eventually prepared by Mr Tony Ayres, former Secretary of the Department of Defence), and spawned the most extensive Senate inquiry ever conducted into the operation of Australia’s refugee and humanitarian processes: Senate Legal and Constitutional References Committee, A Sanctuary Under Review: An Examination of Australia’s Refugee and Humanitarian Processes, June 2000

72 Chen Shi Hai v Minister for Immigration & Multicultural Affairs (2000) 201 CLR 293.


74 Id at 13 (Gleeson CJ).
rather vigourous objections to the two rulings. However, the decisions are in line with the way other courts around the world are dealing with similar cases, as well as with academic writings on the subject of women and refugee status. In this respect, the cases are a sharp reminder that the legal regime for the protection of refugees has an international aspect to which courts all around the world are inevitably paying heed.

This point is also made by the Federal Court decision of *Al Masri*. In that case, it was held that the mandatory detention of a failed asylum seeker pending removal will be unlawful if there is no real likelihood or prospect of removal in the reasonably foreseeable future. In reaching this conclusion, both the trial judge and the Full Federal Court referred to international authorities for guidance in construing legislation providing for administrative detention and the curtailment of fundamental rights and freedoms. The Full Court’s construction of the relevant provisions was also informed by the principle that s196 should be interpreted in conformity with established rules of international law and with Australia’s treaty obligations — the most important of which is Article 9(1) of the International Covenant of Civil and Political Rights.

4. *Refugees and the Rule of Law*

The phenomenon of people taking flight and seeking refuge from natural disaster or man-made strife is as old as human history itself. Neither is there anything new in people responding with hostility and suspicion to the stranger at the door. Why

---

75 See above n 34.
76 For example, the critiques made about the gendered nature of the definition, summarised most recently and comprehensively in Heaven Crawley, *Refugees and Gender: Law and Process* (2001); Kristen Walker, ‘New Uses of the Refugees Convention: Sexuality and Refugee Status’, in Kneebone (ed), above n56, chapter 10.
78 While the issue of constitutional validity was canvassed by the Court, the case was ultimately decided on the basis of statutory construction. After examining the language and context of the relevant provisions, the Court held that Parliament ‘did not turn its attention to the curtailment of the right to liberty in circumstances where detention may be for a period of potentially unlimited duration and possibly even permanent’, but rather had proceeded on the assumption that detention would come to an end through the combined operation of ss196 and 198. See *Minister for Immigration and Multicultural Affairs v Al Masri*, id at 269.
79 This included the so called Hardial Singh principles (named for *R v Governor of Durham Prison; Ex parte Hardial Singh* [1984] 1 WLR 704) — that any power to detain is ‘impliedly limited to a period which is reasonably necessary for that purpose’. The Courts also cited cases following Hardial Singh such as *Liew Kar-Seng v Governor in Council* [1989] 1 HKLR 607; *Re Phan Van Ngo and Ors* [1991] 1 HKLRD 499; *Re Wasfi Mahmood* [1995] Imm AR 311; *Klinsman v Secretary for Security & Anor* [1999] HKLRD (Yrbk) 430; *R (on the application of I) v Secretary of State for the Home Department* [2002] EWCA Civ 888; *R (on the application of Vikasdeep Singh Labana) v The Governor of Campsfield House* [2003] EWHC 410; and *R v Secretary of State for the Home Department; ex parte Saadi* [2002] 4 All ER 785 at 793. The United States Supreme Court decision of *Zadvydas v Davis* 533 US 678 (2001) was also considered.
80 See *Minister for Immigration and Multicultural Affairs v Al Masri*, above n77 at 273–277.
else would so many of the world’s religions seek to correct this human response by
injuncting believers to welcome the stranger, and to give succour to the lonely and
the dispossessed? Asylum seekers who present as boat people seem to occupy a
special place in the pantheon of threats to national security around the world.
Perhaps it is because boats can carry larger numbers of people than it is possible to
smuggle onto the heavily regulated aeroplanes. Perhaps the arrival of unauthorised
boats evoke in our deepest psyche primordial fears of invasion and devastation.

Australians are not unique in reacting with alarm to the phenomenon of boat
people. The so-called ‘Pacific Solution’ and ‘Operation Relex’ were modelled
closely on the US program for ‘interdicting’ boat people from Haiti during the
1980s and 1990s.81 The arrival, or even the threatened arrival, of water-born
asylum seekers has evoked legislative responses in even the most avowedly
‘refugee friendly’ countries.82 At the same time, there is a sense in which
Australia’s response to refugees has been out of the ordinary.

The extraordinary side of the refugee story in Australia has little to do with
either numbers or modes of arrival. It has everything to do with the law or, rather,
with the intersection of Australia’s legal institutions in their dealings with refugees.
There is a sense in which refugees have threatened to bring Australia’s judicial
system to its knees, both literally and juridically. By the end of 2001, the High
Court was becoming overwhelmed by applications for judicial review lodged by
failed refugee claimants.83 At the same time, refugee cases in the High Court have

81 The United States has also enacted different legal regimes for persons detained away from its
mainland territories. For an account of these, see Karen Musalo, Jennifer Moore & Richard
the US legislation are the provisions reducing or removing altogether the right of non-citizens
apprehended in this way to access America’s administrative law systems. See Refugee Act 1980,
Immigration and Nationality Act. Note that moves were also made to restrict the access of illegal
entrants to appeal and judicial review mechanisms. See Illegal Immigration Reform and
Immigrant Responsibility Act 1996; and Karen Musalo, Lauren Gibson, Stephen Knight &
Edward Taylor, ‘The Expedited Removal Study: Report on the First Three Years of
Public Policy 130–145.

82 In 1999, rumours that a boat was making its way to New Zealand spurred the enactment of
immigration detention laws: See Immigration Act 1987 (NZ) s128 (13B) and the discussion in
Roger Haines QC, ‘An Overview of Refugee Law in New Zealand: Background and Current
Issues’, paper presented to the Inaugural Meeting of International Association of Refugee Law
which Canada has responded with unusual defensiveness to boats carrying asylum seekers, see
Judith Kumin, ‘Between Sympathy and Anger: How Open Will Canada’s Door Be?’, in US
ws00_sympathy.htm>.

83 Between 1998–1999 and 1999–2000, applications to the High Court had increased from 65 to
128. By early 2002, there were more than 4000 applicants involved in class actions before the
High Court. See Mary Crock & Ben Saul, Future Seekers: Refugees and the Law in Australia
(2002) at 63.
been at the centre of gargantuan struggles between the government and the judiciary. On the one side is a government intent on stifling the judicial review of refugee decisions on the ground that the determination of protection matters should lie with the executive and with elected politicians, rather than with the unelected judiciary. On the other side are judges imbued with the notion that the courts stand between the individual and administrative tyranny; and that refugee decisions must be made in accordance with the rule of law. In 2003, the battle ceased to be a fight over ‘Protection’ — be it protection of borders or protection of human rights. The fight was all about control, and about the balance of power between Parliament, the executive and the Judiciary within the compact that is the Australian Constitution.

The clash of international and domestic law, with the perceived conflict between ‘rights’ to protection and sovereign powers to control immigration, suggests that refugees will continue to be a source of debate and controversy in Australia as in other parts of the world. The clashes between the executive and judicial arms of government in Australia in refugee cases may have brought little international credit to the country. On occasion, they have also threatened the very fabric of the rule of law in Australia, embodied as this is in the principle that the judiciary alone is vested with the power to make final determinations on questions of law. In *Chu Kheng Lim*, the one point on which the High Court refused to compromise was the constitutionality of a provision that purported to exclude the ability of any court to review the legality of immigration detention. One decade later, the same Court has refused to countenance a broader attempt to oust the curial review of migration decisions.

The Solicitor General, Bennett QC, presented the Minister’s interpretation of the privative clause inserted into the *Migration Act* in October 2001 to the High Court on 4 September 2002. He argued that the effect of the new provisions was to render immune from challenge any interpretation placed on the *Refugee Convention*, provided that the Minister acted in ‘good faith’. Gleeson CJ saw that the Minister’s reading of the clause in question would have enucleated the power of the Court to question or otherwise review the executive’s understanding of substantive refugee law. His Honour brought out this fact by invoking as an example the High Court’s ruling in *Khawar*:

GLEESON CJ: What if, instead of granting a visa to a person who is an alien, he (the Minister) refuses to grant a visa to a person who is a refugee within the meaning of the Refugees Convention?

MR BENNETT: The question then would be whether it was a bona fide exercise of the power.

GLEESON CJ: Suppose the reason he refuses to grant the visa is that he has an erroneous understanding of the meaning of the language of the Refugees Convention?

MR BENNETT: Then, your Honour, it is not challengeable. He is given that power. That is not within the Hickman exception rule.
GLEESON CJ: Is another way of saying that that he is given power to make a conclusive decision as to the question of law involved in the true construction of the Refugees Convention?

MR BENNETT: No, your Honour. That, we would submit, is not a correct characterisation of what he is doing. The effect of that clause, read with the Hickman clause, is that so long as the Minister bona fide attempts to apply what the section requires and so long as one is within the other limitations, then he may make the decision either way and he is empowered to do so.

GLEESON CJ: Take the case that we had recently involving a question of whether women could be a particular social group. Suppose the Minister decides that it is impossible for women to be regarded as a particular social group within the meaning of the Refugees Convention. Does that not amount to the Minister making a conclusive and incontestable decision about a matter of law?

MR BENNETT: No, your Honour, because the Minister is not ultimately deciding a question of law; the Minister is deciding whether to grant a visa and he is making a bona fide effort to apply a criterion which he may apply wrongly.84

Insofar as the Solicitor General was suggesting that the High Court would not be at liberty to make the ruling that it did in *Khawar* under the privative clause regime, one can sense the affront to the Chief Justice. In the result, the High Court was unanimous in its rejection of the Minister’s arguments on the effect of the privative clause. Although it upheld the provisions as constitutional, the Court ruled that the clauses simply had no application in cases where decisions are affected by fundamental legal errors that infringe against ‘inviolable limitations’ or other matters that define the jurisdiction of a decision-maker.85 It is my personal view that the ruling will operate to ensure the continued — intimate — involvement of the courts in determining the correct interpretation of the *Refugee Convention* and Protocol.

The importance of the Courts maintaining their role as interpreters and defenders of the law in the area of refugee protection cannot be overestimated. The Courts may not be able to prevent the political posturing and even manipulation that has characterised the political discourse surrounding refugees and asylum seekers in Australia. However, they are in a unique position to at least moderate the impact of the politicisation process on the refugees themselves. In the area of refugee law, the Australian judiciary can, quite patently, be the last bastion against executive tyranny for the dispossessed and reviled. At risk is life, liberty and the rule of law — not just for the refugee, but for all of us.

---

First Head Revisited:  
A Single Industrial Relations System under the Trade and Commerce Power

DAVID McCANN*

Abstract

The reach of Commonwealth regulation in the industrial sphere has traditionally been constrained by the narrow grant of industrial power in s51(yyyy) of the Constitution. That power allows only for the resolution of interstate disputes by conciliation and arbitration. Increasingly the Commonwealth has sought to extend the applicability and content of its industrial systems by using its other s51 powers. The result with which the Australian industrial system must now cope is a complex hotchpotch of overlapping federal and state mechanisms. This paper argues that the Commonwealth could achieve both broad applicability of its laws, and significant policy-making flexibility, by enacting unitary industrial legislation in pursuance of its s51(i) power to legislate with respect to ‘trade and commerce with foreign nations and among the States’. This power has long been overlooked, despite United States jurisprudence underlying a broad construction. A reconsideration of the untenably narrow understanding of commerce, trenchantly maintained by the Australian cases, is timely given the globalisation of the labour market and the revitalised jurisprudence upon the s92 guarantee of freedom of trade. Commonwealth systems freed from the confines of s51(yyyy) and the complexity caused by reliance on other powers will portend a radical change in the way we conceive of labour relations regulation in this country. The future could be very different under a national industrial relations system based upon the first head.

* BCom (Eco), LLB (Hons) (Melbourne); Articled Clerk, Mallesons Stephen Jaques. The views expressed in this article are those of the author and do not represent the opinions of Mallesons Stephen Jaques. I am grateful to Eloise Dias, Joel Fetter, Martin Joy, Professor Cheryl Saunders and the anonymous referees for comments on earlier drafts. The usual caveat applies.
Introduction

1. The Current Australian Industrial System: Features and Costs 79
   A Dual Track System 80
   B The Exigencies of s51(xxxv) 81
   C Some Conclusions on the Status Quo 82

2. Constitutional Heads of Power 83
   A The Trade and Commerce Power 83
   B The Options under other Powers 85
      (i) A preliminary matter: will the existence of s51(xxxv) restrict the use of other powers to regulate industrial matters? 85
      (ii) Section 51(xxix): The External Affairs 86
      (iii) Section 51(xx): The Corporations Power 87
      (iv) Other Powers 89

3. A Comparative Model: The United States Commerce Clause 90

4. The Australian Trade and Commerce Power 95
   A The Present Understanding of s51(i) Is Too Narrow 95
      (i) The Inter/Intrastate Distinction 95
      (ii) Why the Present Narrow Rule Should be Rejected 97
   B What Construction Should the Trade and Commerce Power Be Given? 101
      (i) Interpretation 101
      (ii) Characterisation 103
      (iii) Future Directions 105

5. Conclusion 105
Introduction

The ‘encumbrances essential to reliance on the Constitution’s industrial power’ have, for most of our nation’s history, severely restricted the utility, efficiency and efficacy of industrial law and policy in Australia. From the earliest possible point in our nation’s history, the Commonwealth has sought to increase its policy choice and the coverage of its industrial systems beyond the strictures of the limited, purposive s51(xxxv). Six failed referenda, one lone referral and a patchwork of provisions under other s51 powers have been the unsatisfactory results.

I argue in this paper that the oft forgotten trade and commerce power is potentially the most helpful power available to the Commonwealth should it seek to legislate for a unitary national industrial relations system.

I assess in Part 1 the features and costs of the present farrago. Dual (state and federal) jurisdictions often overlap. The Commonwealth’s power under s51(xxxv) is exercisable only where there is an interstate industrial dispute, which in turn must be resolved only through a system of conciliation and arbitration. I briefly describe the inefficiencies and uncertainties created by this overlap. A single national scheme of employment standards and dispute resolution mechanisms, administered by the Commonwealth, is presented as the benchmark in this analysis. The comparison indicates the complexity, expense and unfairness that unnecessarily characterise the current system. These impediments result primarily from the narrow terms of the grant of industrial power under s51(xxxv), which still underlies the bulk of federal industrial legislation.

Assessments of the ‘failings’ of the current shared system have been repeatedly made elsewhere. On the other hand, it may well be that in the present political context a shift to a Commonwealth system would eviscerate desirable rights. It could also preclude the policy dynamism that comes with our competitive federalism. My analysis, however, deliberately eschews an overtly political comparison of the existing model and a unitary arrangement. Nor do I examine the entrenched practical and political difficulties that any government would face in trying to push through such a change. This paper instead concentrates on the constitutional contortions presently required and the legal basis for a renewed federal power.

3 See Constitution Alteration (Legislative Powers) 1910 (Cth); Constitution Alteration (Industrial Matters) 1912 (Cth); Constitution Alteration (Railway Disputes) 1912 (Cth); Constitution Alteration (Legislative Powers) 1919 (Cth); Constitution Alteration (Industry and Commerce) 1926 (Cth); Constitution Alteration (Industrial Employment) 1946 (Cth); see also Parliament of Australia, Parliamentary Handbook (‘Referendum Results’) (2003): <http://www.aph.gov.au/library/handbook/referendums> (14 December 2003).
4 Commonwealth Powers (Industrial Relations) Act 1996 (Vic).
5 See, for example, Abbott, below n10.
Part 2 turns to alternative solutions. Labour regulation must now respond to a continuously changing workforce and an employment market that is migratory and essentially national. We can thus expect federal governments in the new millennium to prefer those powers which will provide two touchstones: first, comprehensive coverage, and second, policy design flexibility. Against these familiar criteria, I examine a range of constitutional powers and their capacity to effectively support a broad unitary system.

I then explain why a reconsideration of the trade and commerce power, in particular, is timely. Many commentators have noted that s51(i) has a large degree of unexplored potential. This is especially the case given the High Court has now clarified the meaning of s92 (the guarantee of free interstate trade and commerce), which had previously constrained a broad or plain meaning interpretation of its related power.

In Part 3, I describe the broad interpretation of the commerce clause in the United States Constitution, a clause substantially the same as s51(i). In the US, comprehensive federal labour legislation is based upon this broadly understood clause; it thus provides a useful comparator and model for Australia.

With the United States model in mind, I discuss in Part 4 the traditional interpretation of the Australian trade and commerce power. The High Court’s narrow approach to the power is difficult to sustain theoretically and presents a technical account of commerce that is increasingly incompatible with a modern national economy. I argue that s51(i) is instead properly construed as granting to the Commonwealth a general power to legislate with regard to interstate and certain intrastate commerce. If labour is then characterised as a part of (or affecting) this ‘broad commerce’ then a unitary industrial relations framework, with wide coverage and a high degree of policy flexibility, will be possible under the commerce power. I then briefly indicate the potential for Australian labour law under a newly understood trade and commerce power.

---


1. The Current Australian Industrial System: Features and Costs

Six different industrial jurisdictions continues to make as much sense as having different rail gauges in different states.\(^{10}\)

For the purposes of the following analysis, ‘a unitary industrial system’ is best understood as an arrangement whereby a single level of government has exclusive power to set standards and regulate relationships in the industrial sphere.\(^{11}\) It would preferably involve a single set of accessible Commonwealth laws, a national system of tribunals and a sole administering government. Industrial law-making power in Australia is, instead, divided between six overlapping polities.\(^{12}\) Our unique federal model, enshrined in a stubbornly change-resistant Constitution, involves compromise and duplication in labour market regulation on an almost unprecedented scale.\(^{13}\) Control over the standards and ‘default rules’ in employment relationships is shared. Concurrent, competing jurisdiction is the norm.

This obligatory sharing is due primarily to the limited nature of the Commonwealth’s industrial power:

Section 51. The Parliament shall, subject to this Constitution, have power to make laws for the peace, order, and good government of the Commonwealth with respect to: … (xxxv) conciliation and arbitration for the prevention and settlement of industrial disputes extending beyond the limits of any one State.

The grant is thus limited in two quite perpendicular ways. First, the Commonwealth has power only when there is a dispute, and can only legislate to solve that dispute in certain constitutionally-defined ways.\(^{14}\) Second, the dispute must be interstate before the Commonwealth’s jurisdictional competence is triggered; the States’ power, though subordinate in the event of inconsistency,\(^{15}\) is not excluded per se. Here I examine the consequences of, first, the State/ Commonwealth duplication, and second, the restricted manner in which the Commonwealth may legislate under s51(xxxv).

---


\(^{12}\) The Commonwealth, and each of the states except Victoria, which has referred its industrial power to the Commonwealth under s51(37): Commonwealth Powers (Industrial Relations) Act 1996 (Vic); and see also Stuart Kollmorgen, ‘Towards a Unitary National System of Industrial Relations? Commonwealth Powers (Industrial Relations) Act 1996 (Vic), Workplace Relations and Other Legislation Amendment Act (No 2) 1996 (Cth)’ (1997) 10 AJLL 158.


\(^{14}\) Constitution s51(xxxv).

\(^{15}\) Constitution s109.
A. Dual Track System

The Australian industrial ‘system’ is essentially two parallel systems, federal and state. Both are made up of tribunals, parliaments, government departments, industrial Acts, and so on. Duplication and resource waste is, to some extent, a predictable result when similar types of institutions are doing similar work in similar ways. Efficiencies of scale are unavailable. Each tribunal, for example, must maintain its own registry and conduct its own procedures. The additional costs are often borne by taxpayers, or by participants in the form of less effective service.16

Dual systems, especially dual dispute resolution tracks, also portend inconsistency and inequity in result: it is more likely that there will be unequal treatment of like disputants in different jurisdictions. ‘Rival shops’ enable forum-shopping and the associated uncertainty causes further inefficiency and unfairness.17 Dual federalism also causes certain problems specific to the industrial sphere: delays in flow-on of award changes,18 union demarcation and registration disputes,19 and uncertainty over s109 inconsistency as it applies to awards.20

All these problems have been somewhat minimised by inter-tribunal and inter-governmental cooperation, by mutual recognition of precedents and practice, and by parallel legislative action. To take one example, the NSW Industrial Relations Commission recently ‘passed on’ the increase in the Australian Industrial Relations Commission’s national wage case decision, albeit with a significant delay and to workers who could just as well have benefited from the original decision had they been covered. Institutionally, governments often cooperate by dual appointments mechanisms and other comity provisions; the former NSW Coal Tribunal is perhaps the most illustrative example in this area. Initiatives and work at the Council of Australian Governments, by the Australian judiciary, and in universities, have enhanced the effectiveness of this information-sharing and constructive comparison.

Some competitive federalism theorists argue that six systems each striving for best practice will lead to a more dynamic reforming environment and thus to better outcomes in the long run.21 On the other hand, six poorly resourced and

---

18 Id at 207–209; and see below.
19 Macken, above n17 at 210–14.
administratively burdened systems, each struggling to maintain national coherency in policy and practice, may merely exhibit a drag effect on economic performance and the achievement of workers’ rights. 22

B. The Exigencies of s51(xxxv)

The legislative and institutional contortions demanded by reliance on the conciliation and arbitration power have now become established features of the Australian industrial landscape. Commonwealth industrial legislation is limited in its ends and its means: laws must be for a system of dispute settlement through conciliation and arbitration. 23

Primarily, the Commonwealth cannot use s51(xxxv) to directly regulate industrial relationships or to directly mandate rights and obligations. 25 For example, s51(xxxv) could not support a general minimum wage. 26 Rather, access to the federal industrial jurisdiction requires parties to create a dispute. The logs of claim served in such disputes then define the subsequent arbitration to the ‘ambit’ of the original dispute. 27 This ‘ambit doctrine’ often results in expansive initial claims so as to secure maximum flexibility and minimise later costs.

Further, only parties to the dispute are bound by the award or decisions made. 28 This creates the constant need to ‘rope in’ new parties, and precludes common rule awards 29 that could delineate the rights and obligations of all parties more equally and clearly. As noted above, the dispute must be interstate. Artificial ‘paper


23 Though the power is not purposive: Industrial Relations Act Case, above n2; Re Pacific Coal Pty Ltd, Ex parte Construction, Forestry, Mining and Energy Union (2000) 203 CLR 346 at 412 (Hayne & Gummow JJ) (hereinafter Re Pacific Coal).


26 R v Commonwealth Conciliation and Arbitration Commission; Ex parte Amalgamated Engineering Union (Australian Section) (1967) 118 CLR 219 at 242 (Barwick CJ) at 269 (Windeyer J). Compare Fair Labor Standards Act 1938 (US), 29 USC s206, to which I return below, and the Industrial Relations Amendment Act 1993 (Cth). Note, however, the unexplored ‘prevention’ aspect of the power, which I discuss below at n31 and accompanying text.


28 See Workplace Relations Act 1996 (Cth) s149(1); compare the Act’s transmission of business provisions.

29 Australian Boot Trade Employees’ Federation v Whybrow & Co (1910) 11 CLR 311; R v Kelly; Ex parte Victoria (1950) 81 CLR 64 at 79–82.
disputes’, bringing in unrelated but interstate employers and peripheral employment issues, are perhaps the most egregious illustration of the administrative hula-hoops that are required.  

There are two caveats to this caustic analysis. First, I note the currently underutilised and unexplored ‘prevention’ aspect of the power;31 this may allow common rule awards, for example, though it would probably still require at least a potential dispute. Second, the Workplace Relations Act 1996 (Cth) (hereinafter WR Act) is of course no longer tied exclusively to the industrial power. Indeed, part of the complexity and overlap I have described results from the Commonwealth’s use of other powers, such as external affairs and corporations, which are less prescriptive but cover only limited participants.32

Notwithstanding this, however, the constitutional exigencies of s51(xxxv) cause complexity, cost and lack of coverage33 far beyond that which we might expect under a simple unitary system. The jurisprudence upon the power is nevertheless a mature one. There seems little chance that the judicial interpretation could (or should) be unshackled from the clear terms of what is, at base, a limited, purposive and not plenary power. Despite some inkling in recent decisions,34 it appears unlikely that the Commonwealth could successfully advance a significantly wider view in the High Court.

C. Some Conclusions on the Status Quo

The Australian industrial system is thus conceptually, legally and procedurally complicated and costly.35 This lack of simplicity produces inaccessibility. There is the potential for experienced repeat players (often ‘third party’ unions and employer organisations), and the well resourced, to dominate these processes or cultivate complexity to cement their own advantage.36 This in turn excludes the voices of small, poor and unsure participants, most especially employees and some small business employers.

These problems of overlap, duality and cost have often been explored37 Their depth and intractability is also indicated by the range of governmental action taken in response. Most obviously, cooperation at an administrative and tribunal level has

30 See Williams & Gotting, above n24 at 266; R v Ludeke; Ex parte Queensland Electricity Commission (1985) 159 CLR 178 at 181–2; A–G (Qld) v Riordan (1997) 192 CLR 1 at 39–40 (Kirby J).
31 See Re Federated Storemen and Packers Union of Australia; Ex parte Wooldumpers (Vic) Ltd (1989) 166 CLR 311 at 320–1 (Mason CJ), 327 (Deane J); Williams, above n8 at 47–48 and 83–4.
32 See the analysis and examples below in Parts 2 and 4 and provisions of the WR Act at nn227–230.
33 Wooden, above n6 at 181.
35 A ‘Serbian bog of technicalities’: id at 422 (Kirby J), quoting HB Higgins in Robert Garran, Prosper the Commonwealth (1958) at 225.
36 The current federal government has often made this criticism, in my view quite validly. On the growth of non-union employee relations in recent times see generally John Campling & Paul Gollan, Bargained Out: Negotiating without Unions in Australia (1999).
37 Hancock Report, above n11 at 264–9; Breaking the Gridlock I, above n16.
become a necessary reality. Federal and State comity provisions are extensive. The very fact that the Commonwealth has so often gone beyond the specific industrial power in s51(xxxv) indicates the unsatisfactory nature of that grant.

It is important to note, however, that complete Commonwealth control will not necessarily bring with it simplicity. Nor will it necessarily enhance the relative power of workers, or improve the conditions in which they work. Certain advantages may also be lost under a centralist, unitary federal system: local, on-the-ground dispute resolution may become less flexible and less responsive for example. The dynamic policy improvement that can come from competitive federalism would also be pre-empted — reform would no longer be impelled by the successes of ‘the State next door’.

While I do not here come to a conclusion on the desirability of de- or re-regulation of the Australian labour market, it is the purpose of this paper to explain how Commonwealth legislative capacity may increase upon a reassertion of jurisdiction under s51(1) to opportunities that may open up upon a reassertion of Commonwealth legislative capacity. I have noted above the benefits possible under a unitary system. If the Commonwealth wishes to capture these benefits, it must increasingly look to enlarge its power, if necessary at the expense of the States. I have explained why the deliberately constrained grant in s51(xxxv) requires overly complex and inefficient national industrial legislation. I turn now to examine other heads of power that may potentially be used to expand the Commonwealth’s constitutional competency.

2. Constitutional Heads of Power

A. The Trade and Commerce Power

Given the limitations described above, this section now examines the Commonwealth’s options under several other s51 Constitutional powers. The ‘corporations’ and ‘external affairs’ powers, in particular, have been utilised in the past by a Commonwealth eager to influence the industrial landscape. This article focuses, instead, upon the oft forgotten trade and commerce power. This power could be used to supplement or simplify the jurisdiction asserted under other s51 powers. A new consideration of this first head is warranted for three major reasons.

38 Hancock Report, above n11 at 261.
39 See, for example, the former Industrial Relations Act 1988 (Cth) s67 (regarding State/ Commonwealth Industrial Commissions conferences).
40 See a full review in Macken, above n17 at ch18–19.
41 See, for example, Maritime Industries Act 1929 (Cth) (an early example of reliance on the trade and commerce power); WR Act s5 (concerning access to the AIRC); WR Act pt VIB div 3 (extending the reach of Commonwealth provisions using the corporations power) and WR Act pt IV A div 5 (implementing the ILO Workers with Family Responsibilities Convention 1981 and Workers with Family Responsibilities Recommendation 1981).
42 See, for example, WR Act sch 1A (conditions for Victorian workers).
43 Hancock Report, above n11 at 264.
First, I think the limits upon legislative power with regard to trade and commerce previously set by the High Court are incorrect. This is especially so given the increasing globalising pressures to which Australia is subject. Using a detailed comparative analysis, I hope to show that this power is better understood as sufficiently broad that it could ground a unitary industrial relations system.  

Second, this power provides an effective, if not the most effective, way for the Commonwealth to implement a system that covers both a wide range of workplace issues (policy flexibility) and a large percentage of Australian workers (coverage). Although it is clear that a referral of power to the Commonwealth or a Constitutional amendment would each have conclusive effectiveness, neither referral nor amendment is likely. Other options have been widely canvassed, the most notable of these being a constitutional corporations-based regime. Although industrial legislation based on the external affairs, territories, corporations, or combinations of these and other powers has been validly enacted in the past, there are limitations in both the coverage and policy flexibility available under these heads. For a variety of reasons, progress on reform toward these alternative models seems to have stalled.

And thirdly, contemporary exploration of the potential of s51(i) in regard to Commonwealth industrial regulation is sparse. Despite judicial indication of this potential, a recently enriched and clarified jurisprudence on s92 interstate commerce, and an increasingly professed desire for simplicity and efficiency in the Australian industrial context, few have yet delved into the latent capacity of this head of power. Most labour lawyers do no more than briefly acknowledge

---

46 Constitution s51(xxxvii).
47 Constitution s128. An amendment could simply give the Commonwealth power over ‘industrial relations’.
48 Compare Federal Awards (Uniform System) Act 2003 (Vic) ss50–53.
49 Breaking the Gridlock 1, above n16, Breaking the Gridlock 2, above n22. See further Stewart, above n 27.
50 See below Part 2.B, The Options under other Powers.
52 Cole v Whitfield, above n9 at 360; see below Part 4.A.ii, ‘Why the Present Narrow Rule Should be Rejected’.
53 See, for example, Breaking the Gridlock 1, above n16 and Breaking the Gridlock 2, above n22.
54 See, however, Williams, above n8 at 140–145; Zines, above n8 at 55–79. The last major consideration of the power and labour law was S D Hoptop, 1974. The Commonwealth’s reluctance to urge a wider interpretation, both in constitutional litigation and in policy papers, is curious: see, for example, the cursory discussion Breaking the Gridlock 1, above n16.
the possibilities under s51(i). It seems that a comprehensive reconsideration of the power is both necessary and overdue.

To chart the constitutional ground, I begin in the next section with a consideration of the more popular (and more familiar) elements of the s51 scheme.

B. The Options under other Powers

(i) A Preliminary Matter: Will the Existence of s51(xxxv) Restrict the Use of Other Powers to Regulate Industrial Matters?

It might seem strange and somewhat illicit to seek to avoid the strictures of the industrial power by relying on other powers like external affairs or quarantine. The simple answer to this is that the High Court has, in most circumstances, refused to restrict the meaning of one power because of the existence of another power. 'Secondary' powers, such as the corporations and external affairs powers, have been widely and validly used in the industrial field in the past.

During the darkest days of World War Two, the High Court in *Pidoto v Victoria* questioned whether the defence power, when used for industrial legislation, would be limited by the requirements of s51(xxxv) (requiring, for example, that disputes be interstate). The court held that the limits in s51(xxxv) were expressed not 'by way of exception, but by way of definitional circumscription of an affirmative grant.' Section 51(xiii) on the other hand, which gives the Commonwealth power with respect to 'Banking, other than State banking', positively quarantines certain matters (State banking) from the reach of all Commonwealth legislation, even legislation that would otherwise be valid under another power. It is enough here to note that *Pidoto* stands for the proposition that the s51(xxxv) 'conditions' will not restrict legislative capacity under other powers.

The Commonwealth's constitutional power may also be used to achieve, indirectly, objects that seem to be outside the areas it can regulate directly. For example, intrastate trade can be regulated using the corporations power, and some of the affairs of non-constitutional corporations can be regulated under the trade and commerce power. It is not correct to first seek the best characterisation of the law, and then try to fit that law within the constitutional power which most

---

55 See, for example, Creighton, Ford & Mitchell, above n8 at 332–333.
56 See, for example, *Strickland v Rocla Concrete Pipes* (1971) 124 CLR 468; compare *Bank of New South Wales v Commonwealth* (hereinafter *Bank Nationalisation Case*) (1948) 76 CLR 1 (concerning the 'just terms' guarantee).
57 See further Parts 2.B.ii and 2.B.iii below.
58 (1943) 68 CLR 87.
60 This seems to have been the general approach taken by the majority in the *Industrial Relations Act Case*, above n2; contrast the reasons of Dawson J.
closely meets that characterisation. That would be to invert the proper and logical order of reasoning. Rather, the question is whether the law can be fairly described as a law with respect to one of the s51 heads. And to answer that question, first the powers themselves must be interpreted and understood. In regard to labour law, the general rule has been posited most definitely as follows: ‘A law with respect to overseas trade and commerce is a valid law and does not cease to be valid because it can also be characterised as a law with respect to employment’.

(ii) Section 51(xxix): The External Affairs Power

The external affairs power in s51(xxix) of the Constitution gives the Commonwealth the power to implement international treaties (and some Recommendations) through legislation, so long as such legislation is ‘reasonably capable of being considered appropriate and adapted to implementing the convention’. The *Industrial Relations Reform Act* 1993 (Cth) was the most significant importation of international labour norms into the Australian industrial landscape to date. It appears that the Act was quite deliberately drafted to qualify as ‘implementing legislation’; and indeed the national, unitary employment standards it introduced have, for the most part, withstood constitutional challenge. The Commonwealth was able to directly impose obligations on employers with respect to termination of employment, discrimination, and family leave, for example.

It is necessary to consider two matters in relation to external affairs that are relevant to the broader argument here. First, the Commonwealth has already begun to outline a national, unitary set of employment standards. It has done so not using the industrial power, nor by amendment or referral; instead it has taken an indirect route using another seemingly unrelated head of power. The High Court has

---


64 See below Part 4.B, ‘Characterisation’.

65 *Seaman’s Union v Utah Development Co*, above n51 at 154 (Mason J).

66 *Industrial Relations Act Case*, above n2 at 487–8, 496 (Brennan CJ, Toohey, Gaudron, McHugh & Gummow JJ) approving a test originally developed in *Commonwealth v Tasmania* (hereinafter *Tasmanian Dam Case*) (1983) 158 CLR 1 at 529 (Deane J).

67 See generally *WR Act* pt VIA (Minimum entitlements of employees); *Williams & Gotting*, above n24 at 273–75.

68 *Industrial Relations Act Case*, above n2. Note, however, the effect of a change in government: see generally *Workplace Relations and Other Legislation Amendment Act* 1996 (Cth) (minimum wages repeal); Eichenbaum, above n1 at 68; *Williams & Gotting*, above n24 at n80.

69 *WR Act* pt IVA div 3, implementing (inter alia) the *Termination of Employment Convention* 1982; compare ss170DE(2) & 170EDA(1)(b) which were invalid: *Industrial Relations Act Case*, above n2 at 518.


endorsed these impositions, even despite the indirectness, by interpreting a power of the Commonwealth in a broad and coherent manner. The implication for s51(i) is clear.

Second, it appears from the tests (as primarily developed in the Tasmanian Dam Case,\textsuperscript{73} Chu Kheng Lim,\textsuperscript{74} and the Industrial Relations Act Case\textsuperscript{75}) that an external affairs-based system would be limited in the substantive standards and processes it could establish; policy flexibility may be constrained. Under the ‘proportionality test’ quoted above,\textsuperscript{76} Commonwealth regulation must conform relatively strictly to the underlying international standard. For example, if the Commonwealth wished to eliminate the right to strike, it would be difficult to find a treaty mandating that elimination, and still more difficult to eliminate the right and still be ‘appropriately implementing’ a treaty that ensured the right to strike. Similarly, broad and imprecise treaty standards,\textsuperscript{77} or partial implementation of a treaty,\textsuperscript{78} may cause the Commonwealth regime to fall foul of the validity test for want of ‘appropriateness of implementation’.

Reliance on this power thus shifts influence to the ILO,\textsuperscript{79} and probably reduces the range of policy alternatives available to the federal government. However, selective and innovative use of the raft of ratifiable instruments, and their ambiguities, could give even the deregulator a platter of ‘implementation’ options. Of course, at the present time, the obstacles to greater use of the external affairs power are far more political than legal.

It should be concluded, then, that a Commonwealth regime based on s51(xxix), as presently interpreted, would not provide the legislative flexibility of the more substantively plenary powers, such as ss51(xx) and 51(i). Nor would a scheme based, in essence, on implementation of often complex, rights-based ILO instruments facilitate a programme of labour market deregulation (cf (re)regulation). Although the Commonwealth could potentially bind every employer and employee (coverage would be high), as indeed it has done,\textsuperscript{80} the sacrifice would be less freedom and flexibility for federal policy makers.

(iii) \textit{Section 51(xx): The Corporations Power}

The Corporations power\textsuperscript{81} is typologically different to the external affairs power: it is defined by reference to certain \textit{matters}, but rather to particular commercial \textit{entities}.\textsuperscript{82} The power thus conferred has been described as plenary where trading,
financial or foreign corporations are concerned. In other words, s51(xx) gives the Commonwealth a power which extends to the control and regulations of all activities and relationships of these ‘constitutional corporations’.83

The most recent series of Liberal-National governments has advocated greater reliance upon the corporations power for the purpose of simplifying and nationalising industrial regulation.84 The relatively large number of employees that would fall within the reach of a corporations-based scheme (coverage),85 along with the almost unconstrained ability to regulate where a corporation is concerned (policy flexibility), makes this power attractive on both criteria.86

Legislation with respect to the ‘activities, functions, relationships, and business’ of a foreign, trading or financial corporation will presumptively be within Commonwealth power.87 Industrial relations at corporate enterprises would almost certainly fall within this class of regulatable subject matter;88 employees have a relationship with corporate employers; the functions of a corporation include employment of staff; and (most interestingly in considering s51(ii)) the ‘business’ of a corporation extends to the engagement and marshalling of its workforce. Indeed, the corporations power currently validly grounds, for example, Part VIB Div 3 of the WR Act, dealing with enterprise flexibility agreements.89 It may be the case that a labour regime would need to have a specific or differential application to corporations (as opposed to other employers),90 though the courts have generally been willing to construe even broadly expressed applicability provisions as giving the regime the required specificity of application.

As will be seen below, the latitude the court has given the Parliament in how it regulates corporations should be expected to carry over to any new understanding of how trade and commerce (and therefore industrial relations) may be regulated. If the legislation operates upon a constitutional corporation, no further test of purpose or pre-dominant operation or any other requirement will be imposed.91

This said, a corporations power-based scheme would have two main disadvantages. First, it could only apply to (foreign, trading and financial) corporations, and so only to a portion of the total number of employers. Even

84 See Breaking the Gridlock 1, above n16; Breaking the Gridlock 2, above n22; Abbott, above n10.
85 Breaking the Gridlock 1, above n16 at appendix 4.
87 Re Dingjan, above n83 at 364–5; Re Pacific Coal, above n23 at 375 (Gaudron J).
88 Creighton & Stewart, above n8 at 83–85.
89 Industrial Relations Act Case, above n2 at 539 (Brennan CJ, Toohey, Gaudron, McHugh & Gummow JJ).
90 See Williams & Gotting, above n24 at 270–1; see further Re Pacific Coal, above n23 at 447–8 (Kirby J).
91 Compare below n208 and accompanying text.
though the proportion of employees covered by federal jurisdiction could rise to 85 per cent of non-farm employees,\(^{92}\) at least 15 per cent would remain in state systems or outside both systems. By definition then, the Commonwealth could never cover the field using this power. As I will show below, the trade and commerce power, if properly interpreted, should extend to at least this number of employees, for the vast majority of foreign, trading and financial corporations will be engaged in ‘constitutional commerce’ when that term is realistically understood.

Second, many of those most in need of labour rights and protections would be systematically excluded from a regime of this type. Outworkers, partnerships and independent contractors, for example, will often be working for non-corporate entities, and a sufficiently close connexion to a constitutional corporation will not always be easy to find.\(^{93}\)

Though it is beyond the scope of this paper to determine whether an exclusively s51(i)-based scheme would have statistically broader coverage than an exclusively s51(xx)-based scheme, it is sufficient to note that s51(i) would enable more labour relationships to be brought within the scope of a Commonwealth regime, as it would cover most, if not all, constitutional corporations as well as non-corporate, yet commercial, entities.\(^{94}\) The policy flexibility allowed under the two would be comparable. Of course, if used together, the trade and commerce and corporations grants could perhaps give the Commonwealth its broadest power. There is no reason at law why an industrial regime could not be expressed to operate on corporations and on those in trade and commerce. Indeed, a catch-all such as this is the most prudent legislative policy going forward.

(iv) Other Powers

Innovative use of certain of the Commonwealth’s other powers might also be made, though most of these options involve certain policy rigidities or limited employee coverage. The taxation power could be used to encourage employers or employees to subject themselves to a Commonwealth industrial relations framework, by imposing additional tax burdens upon non-compliers. This power was used to require certain training expenditures by employers in the Training Guarantee Act 1990 (Cth),\(^{95}\) which was held to be a valid piece of ‘taxation’ legislation despite its overriding ‘industrial’ purpose.\(^{96}\) The tax burden mechanism is, however, administratively cumbersome, costly, and again would be a top–down regulatory scheme, rather than a scheme based around employees’ rights.\(^{97}\)

\(^{92}\) Breaking the Gridlock 1, above n16 at appendix 4.

\(^{93}\) See, for example, the finding in Re Dingjan, above n83.

\(^{94}\) See Williams, above n8 at 145.

\(^{95}\) See also Training Guarantee (Administration) Act 1990 (Cth).

\(^{96}\) Northern Suburbs General Cemetery Reserve Trust v Commonwealth (1993) 176 CLR 555 at 569, 579. This same logic has been applied to the relationship between the trade and commerce power and employment regulation: see n65 and accompanying text.

\(^{97}\) See text accompanying n93.
There is a range of other powers that could extend the coverage of a Commonwealth scheme at the margins. Undoubtedly quarantine, postal, telecommunications, customs, lighthouse, and some other specific occupational classes of employees would be bound by a Commonwealth scheme if it were specifically extended to them (and perhaps if not). Similarly, the Constitution gives the national Parliament power over the Commonwealth public service, defence employees, and all employees working in or affecting a territory. These grants are arguably plenary as against the relevant class of worker, though of course the relevant classes are limited in membership.

Thus, in the current deregulatory climate, as the Commonwealth seeks to oust State systems in favour of a single flexible system with broad coverage, perhaps the most appropriate head of power has become s51(i). An examination of the analogous American Commerce Clause is a likely place to begin this analysis.

3. A Comparative Model: The United States Commerce Clause

The ‘commerce clause’ of the United States Constitution empowers the Congress to ‘regulate Commerce with foreign Nations and among the several states and with the Indian Tribes.’ Canada too has ‘eleven functioning labour relations systems’ and its national regime is also based on a commerce power. However, it is valuable to consider the United States situation as a prime comparator for three reasons. First, the text of the clauses are substantially the same; indeed the s51(i) language was copied at Federation from its American predecessor. Second, in Australia and the US, but not in Canada, the Constitution allocates only certain enumerated powers to the federal legislature, preserving residual and general power to the States. Thirdly, the structural and textual similarity between the two federal systems means that United States jurisprudence is a relevant consideration for Australian jurists. And fourthly, the economies of Australia and the United States are, at the opening of the 21st

---

98 See Hancock Report, above n11 at 250–1.
99 Compare, for example, WR Act s5 (maritime and other workers). See below nn228–230.
100 If read down: Acts Interpretation Act 1901 (Cth) s15A; Pidoto v Victoria (1943) 68 CLR 87; compare Workplace Relations and Other Legislation Amendment Act 1996 (Cth) s7A and discussion in Re Pacific Coal, above n23 at 447–8 (Kirby J). These reading down provisions may also apply to a corporations power-based scheme: see text accompanying n90.
101 Constitution s52, perhaps also the s61 executive power/implied nationhood power.
102 Constitution s51(vi).
103 See generally the discussion of Mason J in Airlines of NSW Pty Ltd v New South Wales [No 2] (1965) 113 CLR 54 (hereinafter Second Airlines Case).
105 Arthurs, Carter & Glasbeck, above n13 at 19, and see especially 53–56.
106 Ibid 54–56.
Century, substantially similar. Both are characterised by highly sophisticated and mature consumer markets, strong service sectors and open import/export of commodities and manufactures. For these reasons, it is both foreseeable and desirable that the High Court have regard to the American Commerce Clause decisions.\textsuperscript{109}

It is important to note, however, that there is not a perfect symmetry. As Zines, Mason and Williams all point out in their analyses of the United States jurisprudence,\textsuperscript{110} the Australian Constitution explicitly gives the Commonwealth power over certain ‘commerce-related’ matters, such as banking and corporations.\textsuperscript{111} The United States Commerce Clause, without these adjunct powers, has always needed to do a large amount of ‘constitutional work’ to permit Federal regulation in these areas. In the same way, the American Constitution contains no equivalent of the Australian industrial power.\textsuperscript{112} It is important to remember then that the broad American interpretation described below has, to an extent, been a ‘virtue born of necessity’.\textsuperscript{113}

Constitutionally it seems anomalous that despite the great textual similarity between Art I, s8, cl3, and s51(i), two such disparate interpretations are maintained. Congress first asserted its power to regulate labour relations in the \textit{Railway Labor Act} 1926 (US),\textsuperscript{114} held to be a constitutionally valid protection of interstate commerce in 1930.\textsuperscript{115} Roosevelt’s New Deal took advantage of this early finding with the \textit{National Labor Relations Act} 1935 (US)\textsuperscript{116} (hereinafter \textit{NLRA}) and the \textit{Fair Labour Standards Act} 1938 (US) (hereinafter \textit{FLSA}). These statutes still provide the core federal protection of collective bargaining,\textsuperscript{117} union organisation,\textsuperscript{118} minimum wages,\textsuperscript{119} maximum hours\textsuperscript{120} and other employment standards.\textsuperscript{121} If a United States style interpretation of s51(i) were adopted, we might fairly confidently predict the Australian Parliament would assume a similar

\textsuperscript{109} For the most recent example of the High Court’s use of relevant United States jurisprudence, see \textit{Re Maritime Union of Australia; Ex parte CSL Pacific Shipping Inc} [2003] HCA 43 (7 August 2003) at [36] (discussing wages of seamen on foreign vessels). The impact of this unanimous judgment is discussed further in the text accompanying nn209–215.

\textsuperscript{110} Zines, above n8 at 61; Williams, above n8 at 144; Mason, ‘Role of the Constitutional Court’, above n51 at 15–16.

\textsuperscript{111} See \textit{Constitution} ss51(ii), (v), (xii), (xvii), (xviii), (xx), s52(i), s90 and compare s98, which specifically includes certain things within commerce.

\textsuperscript{112} Though this may not necessarily connote that the US/Australia situation need be different, if the powers under s51 are read as independent: see text accompanying nn56–65 and 96.

\textsuperscript{113} Mason, ‘Role of a Constitutional Court’, above n51 at 15.

\textsuperscript{114} 45 USC ss151–163 and 181–188.

\textsuperscript{115} \textit{Texas & New Orleans Railroad Co v Brotherhood of Railway & Steamship Clerks} (1930) 281 US 548.

\textsuperscript{116} \textit{National Labor Relations Act} 1935 (US) 29 USC ss151–169 (hereinafter \textit{NLRA}).

\textsuperscript{117} \textit{NLRA}, 29 USC ss157–8.

\textsuperscript{118} \textit{NLRA}, 29 USC ss157, 159.

\textsuperscript{119} \textit{FLSA}, 29 USC s206.

\textsuperscript{120} \textit{FLSA}, 29 USC s207.

\textsuperscript{121} See further Alvin L Goodman, \textit{Labor Law and Industrial Relations in the United States of America} (2\textsuperscript{nd} ed, 1984) at 52–4. Note that the \textit{NLRA} has been amended by the \textit{Labor Management Relations Act (Taft–Hartley Act)} 29 USC ss141–181 (1947) and other Acts.
power to legislate directly for labour standards. I concentrate therefore on the contours of the American interpretation.¹²²

The key judicial decision as to the validity of federal labour legislation was made in National Labor Relations Board v Jones & Laughlin Steel Corp.¹²³ The Jones and Laughlin Steel Corporation, a large producer and interstate transporter of manufactured steel, was charged with ‘unfair labor practices’ under the NLRA. The Supreme Court’s reasoning in that landmark case still underpins the scheme of regulation described above. It may be distilled to two relevant postulates:

• First, constitutional ‘commerce’ extends to all those activities that ‘have such a close and substantial relation to interstate commerce that their control is essential or appropriate to protect that commerce from burdens and obstructions’.¹²⁴ This is the case even though the activities may be in the nature of production¹²⁵ or ‘intrastate in character’,¹²⁶ so long as they ‘affect commerce in … a close and intimate fashion’.¹²⁷ Thus, even though the relevant employees were solely engaged in the (intrastate) manufacturing of steel, the Congress could empower them with certain labour rights because their employment affected the relevant commerce.

• Second, the Court confirmed its earlier finding¹²⁸ that ‘the [commerce] power is plenary and may be exerted to protect interstate commerce ‘no matter what the source of the dangers which threaten it’’,¹²⁹ including the ‘danger’ of labour stoppages, which ‘have a most serious effect upon interstate commerce’.¹³⁰ The NLRA’s regulation of bargaining and union organisation were thus valid, for ‘it is idle to say that the effect would be indirect or remote’.¹³¹

¹²² For a wider discussion of United States labour law, see especially William Gould IV, A Primer on American Labor Law (3rd ed, 1993).
¹²⁴ Jones & Laughlin, above n123 at 37.
¹²⁵ Id at 34–35 and see conclusion at 43.
¹²⁶ Ibid.
¹²⁷ Id at 32. See also at 37: ‘close and intimate relation to interstate commerce’. Jones & Laughlin, above n123, thus rejected the narrow understandings of ‘affecting commerce’ in Carter v Carter Coal Co (1936) 298 US 238, and while still paying lip-service to conservation of the federal balance, this was the case. See further Zines, above n8 at 55–60.
¹²⁸ In the Second Employers’ Liability Cases 223 US 1 at 51.
¹²⁹ Jones & Laughlin, above n123, approving the rule in the Second Employers’ Liability Cases 223 US 1.
¹³⁰ Jones & Laughlin, above n123 at 41.
¹³¹ Ibid.
It is useful to say one other thing about the reasoning of the unanimous Court in *Jones & Laughlin*. It acknowledged the restriction ‘upon the federal power which inheres in the constitutional grant’. If legislation were allowed to ‘destroy the distinction’ between internal and interstate commerce, then it would be invalid, for it would simultaneously destroy the nation’s ‘dual federalism’. These fears and limitations, it will be seen, currently decide the matter in the Australian jurisprudence. But, as the two postulates above show, the court did not destroy this distinction at law in *Jones & Laughlin*. Rather, it maintained the duality, while simultaneously recognising that legislation under the power might necessarily blur or cross the line if regulation of intrastate commerce was ‘essential or appropriate’ to regulate interstate commerce. I argue that such an approach is suitable in the Australian situation: an implied distinction between the two types of commerce should not obliterate what is, at its heart, an explicit affirmative grant of power.

A series of subsequent cases confirmed and broadened the reasoning in *Jones & Laughlin*. In *US v Darby* a unanimous Supreme Court held that the FLSA validly applied to a manufacturer of lumber who was only partly engaged in interstate trade. The case applied the test developed in *Jones & Laughlin* to a small largely intrastate producer, and in doing so gave an ‘expansive definition to Congressional authority under the Commerce Clause.’ In *Wickard v Filburn* the interpretation reached its zenith. There, the Court held that, even if a farmer is producing wheat for his own consumption and with no intention of selling it, that wheat is ‘in commerce’ because it affects the broader wheat market. The principle was expressed clearly by Jackson J: activities will affect commerce, and thus be within the power of Congress to regulate, if they exert a ‘substantial economic effect’ upon interstate or overseas commerce. This substantial effect

---

132 Id at 30.
133 Ibid.
134 See below Part 4.A.i, ‘The Inter/Intrastate Distinction’.
135 United States v Darby (1941) 312 US 100.
137 *Wickard v Filburn* (1942) 317 US 111.
138 The extent of the effect of Mr Filburn’s non-participation in the broader wheat market is assessed against a (hypothetical) market in which Mr Filburn participated. That his non-participation had some effect upon the market, through either reducing demand or reducing supply, must certainly be true economically. The interesting question is whether the effect is substantial.
139 *Wickard v Filburn*, above n137 at 124–5 (Jackson J). It is true, as Professor Zines notes, that United States courts accept a wide range of evidence relating to economic and social matters, and that the American cases thus often proceed on the basis that the judges ‘can determine and [are] capable of determining the economic nexus between the local activity and interstate commerce’ (Zines, above n8 at 60). Australian courts, for example in relation to the assessments of ‘substantial lessening of competition’ under s45, s50 and parts of s47 of the *Trade Practices Act* 1974 (Cth), are also now routinely charged with this type of ‘pure’ economic assessment. It would be wrong to assume then that the courts would now refuse (or be unable) to make such assessments if a test of economic effect is required or desirable under a broader interpretation of s51(i).
test was applied in later cases and is now a firmly established part of United States law. It is also a legally defensible and economically accurate way to understand the true nature of the grant in s51(i) of the Australian Constitution.

The recent case law has evinced a shift away from a broad construction of the Commerce Clause, and a reassertion of the judiciary’s right to review and invalidate legislation made under that clause. In both US v Lopez and US v Morrison, over strong dissents, the majority invalidated certain federal laws that purported to regulate conduct occurring solely within a state. However, all the justices but one continued to apply the substantial effect test. Williams argues that the narrower understanding of the Commerce Clause may encourage Australian jurists to refer to Supreme Court cases. It does indeed seem more likely that ‘contemporary High Court judges looking to escape the rigidities of past jurisprudence upon s51(i)’, will find an easier analogy in the reduced though still persuasive United States jurisprudence.

It is useful to note one other aspect of the American legal landscape. Unlike their Australian equivalent, both the NLRA and FLSA begin with ‘congressional findings’ (of fact) and ‘declarations of policy’. These recite, in broad terms, the effects of labour disputes upon commerce, and then assert a clear link between the Clause and the industrial legislation by declaring it to be ‘the policy of the Act’ to prevent these disputes so as to protect commerce. The courts in the United States generally defer to these legislative findings (and in so doing uphold legislation as sufficiently connected to commerce), so long as they have a ‘rational basis’ and the implementation mechanism selected is ‘reasonable and appropriate’. Though there is not the same doctrine of ‘judicial deference’ in Australia, the Parliament might consider inclusion of similar findings of fact as a ‘causative hook’ if attempting broad reform using s51(i).

Thus, the United States labour market is essentially now a national and nationally regulated one. Directly determined and legislated minimum standards and dispute resolution mechanisms apply under a broadly understood
interstate commerce power. Overlap and the associated costs are minimised, employment rights are clear and accessible, and the Congress can directly determine and protect the rights and obligations of participants in the market.

4. **The Australian Trade and Commerce Power**

*Under this power as it currently stands, the Commonwealth cannot regulate purely intra-state trade and commerce .... [However, with] the advent of rapid transportation and communication, and the development of modern technology, trade within each State has become intricately connected with inter-state and overseas trade. And the nationalisation of the economy has necessarily expanded the concept of inter-state trade to embrace activities and transactions that formerly had local significance only. These developments might conceivably justify a re-interpretation of the trade and commerce power, the exiting interpretation of which may be anchored in the artifices of legal formalism.*

In this section I argue that the Commonwealth already has a power that, when properly understood, provides a basis upon which to regulate those employment relationships in or affecting interstate trade and commerce. I begin with an examination of the current, narrower judicial view of that clause and then argue that that view is mistaken. In the second part I describe a positive alternative interpretation based on the United States model. I then indicate what a workplace relations system under a more broadly understood power could look like.

A. **The Present Understanding of s51(i) is Too Narrow**

1. **The Inter/Intrastate Distinction**

In a series of cases culminating in *Attorney-General (WA) v Australian National Airlines Commission*, the grant of power to the Commonwealth under s51(i) has been understood in a narrow and increasingly unsustainable way. The primary reason for this interpretation has been the court’s desire to preserve a rigid separation between intrastate and interstate (and foreign) commerce. This has barely masked a desire to protect a notion of dual federalism and an often unexplained antipathy to United States authorities, particularly the ‘commingling doctrine’.

Judges have relied primarily upon the terms of s51(i) to derive this distinction:

---

152 *Attorney-General (WA) v Australian National Airlines Commission*, (1976) 138 CLR 492 (hereinafter *Third Airlines Case*).
153 The judicial history is complex: see especially Zines, above n8 at 60–74.
154 See, for example, *Second Airlines Case*, above n103 at 115 (Kitto J); see below text accompanying nn175–181.
155 See the discussion in *R v Burgess; Ex parte Henry* (1936) 55 CLR 608 of *Houston, East and West Texas Railway Co v United States*, (1914) 234 US 342; see also Williams, above n8 at 143.
The Parliament shall, subject to this Constitution, have power to make laws for
the peace, order, and good government of the Commonwealth with respect to: —
(i) Trade and commerce with other countries, and among the States.

The distinction is thus justified by a negative implication: because intrastate trade
is not mentioned. This ‘controversial distinction’ between intra- and
interstate trade was attributed undeserved importance from an early stage by Dixon
J (in *R v Burgess* and *Wragg v NSW*) and his ‘early view proved to be
influential’. So as to proactively protect a sphere of State power, a type of
implied constitutional prohibition was developed and continued: the
Commonwealth could not legislate in any way that would ‘obliterate’ the clear
‘distinction’ that the Constitution supposedly draws. Dixon wrote:

The distinction which is drawn between inter-State trade and the domestic trade
of a State … may well be considered artificial and unsuitable to modern times.
But it is a distinction adopted by the Constitution and it must be observed
however much interdependence may now exist between the two divisions of trade
and commerce which the Constitution thus distinguishes.

Intrastate commerce was thus effectively quarantined. If intra and interstate trade
were commingled, regulation of that mixed activity would be beyond s51(i): an
unusual result as against the United States authorities and interpretations of other
s51 powers. The better view was put by Murphy J (dissenting) in the *Third Airlines
Case*:

Sections s51(i) and 92 make a distinction between trade and commerce among the
States and that which is not, but do not make trade and commerce among the
states and intrastate trade and commerce mutually exclusive.

Dixon J stated that the clause contained an ‘express limitation’ preventing
Commonwealth interference in intrastate trade. Clearly the limitation is not
express. Intrastate commerce is not mentioned in the text of s51, nor anywhere else
in the Constitution. As Gleeson CJ has recently said, it is one thing to recognise

---

156 Section 98 of the *Constitution* similarly fails to mention intrastate trade directly. Section 98
specifically includes in the trade and commerce power of the Commonwealth the power to
regulate certain shipping, navigation and state owned railways.
157 Mason, ‘Role of a Constitutional Court’, above n51 at 17.
158 *R v Burgess; Ex parte Henry* (1936) 55 CLR 608.
160 Zines, above n8 at 68.
161 *Wragg v NSW* (1953) 88 CLR 353 at 386 (Dixon CJ, with whom McTiernan, Williams, Fullagar
& Kitto JJ agreed).
162 Id at 385–6.
163 *Third Airlines Case*, above n152 at 530.
164 *R v Burgess; Ex parte Henry* (1936) 55 CLR 608 at 672.
165 See *Australian Coarse Grains Pool v Barley Marketing Board (Qld)* (1985) 157 CLR 605 at 635
(Mason J); Zines, above n8 at 77.
the limits of a s51 power; it is quite another to find in that limited grant ‘some negative implication amounting to a prohibition.’ It is here that my main argument sits: the implied distinction between the two types of commerce should not obliterate what is, at its heart, an explicit affirmative grant of power.

Both the Second and Third Airlines Cases, while affirming the Dixonian distinction, actually found that Commonwealth regulation of intrastate flights was valid. The majority in Second Airlines developed a complicated physical/economic effects dichotomy to underpin its decision: intrastate trade which physically affected interstate trade could be regulated, while effects that were economic would not be enough to give the Commonwealth power. This distinction has been roundly criticised, and it would be difficult I think to maintain before a modern court. It seems clear on its face that the exclusion of economic considerations from the construction of a commerce power, so as to buttress an implied prohibition of a type discredited in Engineers, should invite some reconsideration of the rigid distinction.

It is important to point out, by way of contrast, that there are several express constitutional prohibitions relating to commerce: for example, Commonwealth commercial regulation may not ‘give preference to one State or any part thereof over another State or part thereof’, and nor may the Commonwealth ‘by any law or regulation of trade or commerce, abridge the right of a State or of the residents therein to the reasonable use of the waters of rivers for conservation or irrigation.’ These explicit reductions of the scope of the Commonwealth’s legislative prerogative would almost certainly continue, no matter the breadth of the reading given to s51(i).

(ii) Why the Present Narrow Rule Should be Rejected

This implied prohibition should no longer be a part of Australian law; the technical distinction between inter- and intrastate commerce should no longer constrain the clear terms of s51(i). Sir Anthony Mason has rejected the narrow view of s51(i) as

---

166 Re Pacific Coal, above n34 at 360 (Gleeson CJ).
167 Second Airlines Case, above n103; Third Airlines Case, above n152.
168 See above n154 at 115 (Kitto J) and 88 (Barwick CJ). Applied by Stephen J (with whom Barwick CJ & Gibbs J relevantly agreed) in the Third Airlines Case, above n152 at 510.
169 Third Airlines Case, above n152 at 530–1 (Murphy J); Zines, above n8 at 75–77; Leslie Zines, ‘Engineers and the “Federal Balance”’ in Michael Coper & George Williams (eds), How Many Cheers for Engineers? (1997) 81 at 86; Williams, above n8 at 142. Although Mason did not strictly need to consider s51(i), he did comment that ‘No distinction can or should be drawn between what is physically necessary and what is economically necessary’: Third Airlines Case, above n152 at 523.
170 Third Airlines Case, above n152 at 530–1 (Murphy J).
171 Amalgamated Society of Engineers v Adelaide Steamship Co Ltd (1920) 28 CLR 129 (hereinafter Engineers Case). See below nn175–182 and accompanying text.
172 Constitution s99. For example, the Commonwealth could not validly require a higher minimum wage in Tasmania alone.
173 Constitution s100.
‘restrictive and formalistic’ and has advocated instead ‘a more policy-oriented approach to constitutional interpretation’. There are four main reasons why Sir Anthony is correct. First, the present rule requires a version of the reserved State powers doctrine rejected since the Engineers Case. Second, the restrictive interpretative methodology applied to s51(i) is anomalous in the context of other Commonwealth powers. Third, the narrow Airlines rule deprives the s51(i) grant of the bulk of its essential content. And fourth, the cogent and unanimous decision in Cole v Whitfield means that the closely related s92, for the first time in constitutional history, no longer throws a long shadow over s51(i). I will discuss each of these reasons in turn.

A concern to ‘protect’ Australia’s ‘dual federalism’, by preserving certain powers to the States, often lies behind the judicial resistance to a broader interpretation of s51(i). Up until 1920, the High Court read the powers in s51 subject to a theory of reserved state power: ‘grants of power to the Commonwealth were interpreted so as not to encroach on matters traditionally within the legislative realm of the States.’ This doctrine was clearly rejected by the court in the Engineers’ Case, a rejection that many subsequent cases have confirmed. Since that time the powers in s51 have generally been read according to their terms, an interpretative methodology most often favourable to the Commonwealth. In the Third Airlines Case Murphy J concluded that the ‘maintenance of the supposed division [between inter and intrastate commerce] keeps the pre-Engineers ghosts walking.’ Zines similarly concludes that the distinction is ‘lacking in logic unless one adopts a doctrine of reserved powers.’ Of course, a shift to a broad, United States-style understanding of the commerce clause need not dissolve the jurisprudential differentiation between intra and interstate commerce. Although the development I am advocating may well, in fact, allow the Commonwealth to reach far further into the internal commerce of a state, this should not be because the legal definitions are ‘obliterated’; rather, the Court should recognise that the connectedness of commerce and thus a broad interpretation of s51(i) are ‘the inevitable consequence of national economic integration’.

---

174 Mason, ‘Role of a Constitutional Court’, above n51 at 22.
175 See, for example, Second Airlines Case, above n103 at 115 (Kitto J); Compare Dawson J in the Industrial Relations Act Case, above n2. See also Brian Galligan, A Federal Republic: Australia’s Constitutional System of Government (1995) 188.
176 Williams, above n8 at 128–9.
177 See, for example, Uther v Federal Commissioner of Taxation (1947) 74 CLR 508 at 529 (Dixon J); Strickland v Rocla Concrete Pipes, above n56 at 485 (Barwick CJ); Tasmanian Dam Case, above n66 at 128 (Mason J).
178 See, for example, Keven Booker & Arthur Glass, ‘What Makes the Engineers Case a Classic?’ in Michael Coper & George Williams, How Many Cheers for Engineers? (1997) at 63–65. Professor Galligan regrets this shift due to its consequences for the federal balance: see Brian Galligan, above n175 at 188.
179 Third Airlines Case, above n152 at 530 (Murphy J).
180 Id at 530.
181 Zines, above n8 at 77.
182 Geoffrey Sawer, Australian Federalism in the Courts (1967) at 206. See also Mason, ‘The Australian Constitution’, above n51 at 756.
The second reason why the restricted interpretation of s51(i) should be revised is that it is anomalous against the ‘plain meaning’ approach adopted for s51 powers. This approach construes each of the heads of power according to its terms, and resists the imposition of limits derived from the more narrow expression of overlapping powers. For example, it would be curious to find that ‘trade and commerce’ did not extend to the regulation of telecommunications merely because of the existence of s51(v). The better view, I think, is that:

[i]n the interpretation of the Constitution it is a mistake to assume too readily that the presence of a particular head of power reflects an assumption on the part of the framers that the subject matter was not included in another head of power. A competing hypothesis of at least equal strength is that the former head of power was expressed as it was because doubts were entertained as to the scope of the latter.

In the case of the external affairs and perhaps other powers, implicit limitations have been expressly rejected in favour of reading the grant ‘with all the generality that the words will admit’. This approach has been widely accepted as a ‘principle of constitutional interpretation.’ In the Tasmanian Dam Case, Brennan J held that the application of this broad ‘canon of construction to the affirmative grants of paramount legislative powers gives the Constitution a dynamic force which is incompatible with a static constitutional balance.’ The narrow construction of the trade and commerce power denies this dynamic force in an outmoded attempt to maintain a static constitutional balance.

It is worth remembering, too, that trade and commerce is the first of a long list. Though it appears that the powers are somewhat randomly ordered, the primacy of ‘trade and commerce’ reflects the importance of the Federation goal of economic fluidity and integration. The drafting may thus reflect, at least partly, the significant place that the founding fathers envisaged for s51(i).

Third, limiting the Commonwealth’s power to only interstate trade is a technical legal idea that is increasingly difficult to apply coherently to the facts. Lines between intrastate and other trade are difficult to draw: exports and interstate commerce are now heavily entwined with and affected by intrastate commerce.
Corporations and traders are increasingly using sub-contracting and other techniques that demand greater reliance upon others; others that often operate outside of their state. The tools of business are, likewise, increasing the interconnectedness of even the smallest sole traders. If my local corner shop accepts payment for the nationally produced morning paper using an EFTPOS system maintained on a computer in India, it seems patently artificial to attempt to maintain that that paper purchase is outside constitutional commerce.

All but the hardest originalist would concede that this factual change must have implications for the interpretation of the Constitution:

> [t]he complexity of modern commercial, economic, social and political activities increases the connections between particular aspects of those activities and the heads of Commonwealth power and carries an expanding range of those activities into the sphere of Commonwealth legislative competence.\(^{190}\)

As the nature of commerce has changed so has the traditional approach become more difficult to maintain.\(^{191}\) The Constitution can no longer live for horse shoes when everyone is driving Toyotas.

Fourth, the relationship between s51(i) and s92 has been a major factor in all past interpretations of the power.\(^{192}\) Section 92, which provides for the absolute freedom of interstate trade, commerce and intercourse, has been one of the most litigated provisions in the Australian Constitution. In *Cole v Whitfield*,\(^ {193}\) a unanimous High Court determined that this provision only prevented imposition of ‘discriminatory burdens of a protectionist kind’. The decision overturned a ‘fractured jurisprudence’\(^{194}\) and represents a ‘new and sensible beginning’ in the area.\(^{195}\) The Court itself observed that s51(i) and s92 now ‘sit more easily together’.\(^{196}\) Williams concludes that most national legislation of general operation under s51(i), including relevantly a national industrial scheme of industrial regulation, will not fall afoul of s92.\(^ {197}\) This new context means two things for the interpretation of s51(i). First, it gives the Court good reason to re-examine the rules in the *Airlines Cases* and somewhat pre-empts any stare decisis problem. Second, the simpler discriminatory burdens test allows and even encourages a broader construal of s51(i). Previously, the understanding of s51(i) was narrowed and the power constrained so as to ‘protect’ various strained and

---

\(^{190}\) *Tasmanian Dam Case*, above n66 at 221 (Brennan J). This statement was explicitly affirmed by the majority in the *Industrial Relations Act Case*, above n2 at 485.

\(^{191}\) See further Mason, ‘The Australian Constitution’, above n51 at 756.

\(^{192}\) See the comments of Murphy J in the *Second Airlines Case*, above n103.


\(^{194}\) Williams, above n8 at 130.


\(^{196}\) *Cole v Whitfield*, above n9 at 398.

\(^{197}\) Williams, above n8 at 130–131.
complex understandings of s92.198 That need no longer be the case; the new Cole v Whitfield beginning means that s51(i) need no longer be ‘deprived of its essential content’.199

It is my argument then that the strict separation and ‘implied immunity’ is unsustainable both in law and in the present day factual context. A better approach would be to acknowledge, as the Supreme Court of the United States has done, that the power over interstate commerce can extend to some intrastate commerce and production.

The question then becomes: what should be the connection required, between interstate and intrastate commerce, such that power over the former could permit regulation of the latter?

B. What Construction Should the Trade and Commerce Power Be Given?

It falls now to postulate a positive position, first, on how ‘trade and commerce’ should be interpreted and defined: and second, on how a law regulating industrial matters might be characterised such that it came within a s51(i) jurisdiction. The interpretation/characterisation approach is a useful reasoning tool, though the two are not strictly separable and at points the distinction between them, like the distinction between inter and intrastate commerce, ‘tends to break down’.200 I end this section by broadly forecasting some of the possible effects a revitalised first head could have in the industrial field.

(i) Interpretation

Firstly, it will fall to the courts to interpret the Constitution: what precisely does ‘trade and commerce’ mean? Although ‘it is hard to describe the content of the trade and commerce power without referring to examples of laws which are inside or outside that power’,201 the creation of a general rule will be important for the coherent development of a s51(i) jurisprudence. So as to more clearly indicate the way forward, the following discussion is structured around statements of a possible test — What will be the connection required between constitutional commerce and other commerce?

It is not within the scope of this paper to examine which test is most appropriate, or how the High Court is most likely to find in a future case. It is sufficient to point out that the interpretation ought to focus on what is necessary or appropriate to make the grant of power over constitutional commerce effective. That is the slab of stone which the courts should attack with their judicial chisels. The precise features of the statute so produced will and should emerge gradually, and will depend on the facts and frequency of cases that provide ‘suitable vehicles for agitation of the question.’ A number of possible formulations are here to indicate the contours of the slab. The Commonwealth could be permitted to legislate with respect to:

198 See, for example, the individual rights conception in the Bank Nationalisation Case, above n56.
199 Cole v Whitfield, above n9 at 399.
200 Blackshield & Williams, above n59 at 653.
201 Booker, Glass & Watt, above n63 at 48.
• commerce that relates to constitutional commerce;
• commerce that affects constitutional commerce;
• commerce ‘directly affecting’ constitutional commerce;
• commerce affecting constitutional commerce in a ‘close and intimate’ way;
• commerce that it is ‘essential’ or ‘appropriate’ to regulate so as to regulate interstate commerce;
• commerce the regulation of which is ‘incidental’ or ‘ancillary’ to the regulation of constitutional commerce; or
• commerce which has a sufficient temporal, physical or economic nexus to constitutional commerce.

No matter which wording is chosen, the court ought to frame its test in terms of a connection to constitutional commerce; I do not think it could coherently define constitutional commerce such that it included all intrastate trade and production. Most probably, the Commonwealth would attempt to broadly assert its jurisdiction: the relevant law would purport to bind all employees and employers ‘in’ that commerce which ‘relates to’ constitutional commerce. Alternatively, the more direct route may be taken: the relevant law could purport to bind all employees and employers who ‘affect constitutional commerce’.

Of course, however, a scheme limited to narrow constitutional commerce would suffer many of the same disadvantages as a corporations power-based scheme. It would not, however, involve any purposive element as is sometimes seen for the external affairs power. Section 51(i) is not restricted to laws for the ‘protection or development’ of trade and commerce. Rather, ‘it is a power to make laws with respect to overseas trade and commerce and, subject only to express limitation, it extends to forbidding … commerce itself or anything occurring in or directly affecting such trade and commerce.’

---

202 See NLRA 29 USC s152(7) (definition of ‘affecting commerce’). The NLRA defines ‘affecting commerce’ as ‘in commerce, or burdening or obstructing commerce or the free flow of commerce, or having led or tending to lead to a labor dispute burdening or obstructing commerce or the free flow of commerce.’ See also Ji Case Co v National Labor Relations Board (1944) 321 US 332 at 450.

203 Menzies J in Redfern v Dunlop Rubber Australia Ltd (1964) 110 CLR 194 at 219-220; compare Fullagar J in O’Sullivan v Noarlunga Meat Ltd (1954) 92 CLR 565 at 598 (‘with respect to’ connotes all matters which may affect the subject matter of the power).

204 The ‘Jones & Laughlin American test’ – see above n127.

205 Constitution s51(xxxix); Grammall v Marrickville Margarine Pty Ltd (1955) 93 CLR 55 at 77 (Dixon CJ); see also Peter Hanks, Australian Constitutional Law: Materials and Commentary (5th ed, 1994) at 613.

206 See text accompanying n182: the Court should not dissolve the two branches of intra- and interstate commerce, but rather acknowledge that they are interconnected. It is the separation of the two that requires the fictitious formalism, not their ontological separability.

207 A narrow construction of commerce could, for example, require a Commonwealth scheme to touch only those directly employed by businesses directly selling the majority of their product interstate.

I noted earlier that the High Court has not examined the trade and commerce power for a great many years. That was until a very recent decision from a unanimous seven member bench in *Re Maritime Union of Australia; Ex parte CSL Pacific Shipping Inc.* 209 In a short and unambiguous analysis, the Court declined to invalidate or read down s5(3) of the *WR Act*, which extended the jurisdiction of the Australian Industrial Relations Commission to industrial ‘issues’ between ‘employers and maritime employees’ engaged in constitutional commerce. 210 That declination, by itself, seems not to break any great amount of new ground. However, in the course of the judgment, the Court held that:

[a] ship journeying for reward is in commerce; those who co-operate in the journeying of the ship are in commerce and the wages of those persons and the conditions of their employment relate to that commerce.211

The Court’s use of the words ‘relate to’212 is significant. If the connection required between constitutional commerce and the employment conditions regulated is one of mere ‘relation’, that burden would be less than difficult to discharge for the majority of Australian workers. Similarly, ‘those who co-operate’ is a broad formulation, and implies that a person will be in commerce even if that person is not directly so engaged. The Court also made a point of referring approvingly to a United States authority on the United States Commerce Clause,213 indicating that such authority is at least relevant in construing our own s51(i). 214 Despite these good omens, it would be unwise to extrapolate too far, especially given the law at issue was itself restricted to maritime workers engaged in constitutional commerce narrowly defined, and given the ‘neatness’ of the facts: the relevant employees were all foreign nationals engaged as crew of a journeying ship.215

(ii) Characterisation

The second question would be one of characterisation. This inquiry reads the law and its effects so as to ascribe to that law a character which brings it within, or takes it outside, the grant of power. In this way, characterisation looks to the connection required between the matters regulated (for example, employment conditions) and

210 *WR Act* s5(3)(b).
211 *Maritime Union of Australia*, above n209 at [36]. [Emphasis added.]
212 Id at [30] and [36].
213 See id at [37], citing *Patterson v Bark Eudora* (1903) 190 US 169; *Strathearn Steamship Co Ltd v Dillon* (1920) 252 US 348; *Benz v Compania Naviera Hidalgo* (1957) 353 US 138; *McCulloch v Sociedad Nacional* (1963) 372 US 10; all of which relate to the conditions of maritime workers on vessels engaged in commerce.
214 See text accompanying n107–109 above (where I argue that the US jurisprudence is an appropriate interpretative model).
215 See *WR Act* s5(3)(b)(i)–(iii).
the *interstate or overseas commerce*, however defined. Would a law that dealt with employment relationships be sufficiently associated with constitutional ‘trade and commerce’ that it would be a valid law under the terms of s51(i)? This question is difficult to deal with in abstract, and thus a matter which is best left largely aside for the moment, in the absence of test industrial legislation.

However, there are certain general statements of principle that will guide judicial officers in their characterisation of an impugned law. First, the character of the law ‘must be determined by reference to the rights, powers, liabilities, duties and privileges which it creates’. Second, the law must be able fairly to be characterised as being ‘on’ the relevant subject, here constitutional commerce. ‘The court does not demand to be satisfied that the challenged law is a law “with respect to” the relevant topic, but only that it can fairly be described in that way.’ This approach is both broad and ‘permissive’. Once a connection is found to exist between the law and the head of power, ‘the law will be “with respect to” that head of power unless the connection is “so insubstantial, tenuous or distant” that it cannot sensibly be described as a law “with respect to” that head of power’. If provisions were included that confined the regime’s application to employees ‘engaged in’ or ‘cooperating in’ (or whose employment ‘related to’) constitutional commerce, then the law would probably satisfy this requirement. Similarly broad provisions are found in the *NLRA* and *FLSA*.

There are quite definite signs that labour regulation will generally come within trade and commerce as judicially understood. First, in *Australian Steamships Limited v Malcolm*, it was held that Parliament can validly regulate the conduct of persons employed in activities that are part of overseas and interstate trade and commerce. This law is ‘well settled’. Second, the commerce clause in the United States has long grounded extensive federal labour standards and union organisation mechanisms, even though these matters do not ‘seem’ commercial in nature. Third, the power has been utilised before by the Australian Parliament, especially with regard to waterside workers’ and flight crew officers’ access to

---

216 See *Grain Pool of Western Australia v The Commonwealth* (2000) 202 CLR 479 at 492; affirmed by the recent unanimous decision in *Maritime Union of Australia*, above n209 at [35].
217 Ibid.
218 Blackshield & Williams, above n59 at 652 (emphasis in original).
219 Ibid.
220 *Maritime Union of Australia*, above n209 at [35], citing *Re Dingjan*, above n83 at 369. [Emphasis added.]
221 29 USC s152(7) (definition of ‘affecting commerce’).
222 29 USC s203(b) (definition of ‘commerce’) and, for example, 29 USC s206(a)(2) (minimum wage).
223 (1914) 19 CLR 298.
224 Id at 329–330.
225 *Maritime Union of Australia*, above n209 at [36].
227 See *WR Act s4(1).*
the AIRC,\textsuperscript{228} unfair dismissal remedies,\textsuperscript{229} and agreement certification.\textsuperscript{230} So long as the commerce is within s51(i), courts have commonly held that labour matters affecting that commerce are also covered.\textsuperscript{231}

Thus, we might surmise that the width of the interpretation would determine the coverage of any scheme (in other words, how many of those working in, say, purely intrastate commerce would be bound by the Commonwealth law). The best view on characterisation is that a scheme would continue to be valid, no matter how wide the judicial understanding of constitutional commerce, so long as it were expressed to apply only to a defined group — for example, those whose employment ‘relates to constitutional commerce’.\textsuperscript{232}

(iii) Future Directions

Depending on the understanding of what constitutional commerce is, and which types of matters sufficiently affect it, many new types of employment relationships could be brought within the scope of the Workplace Relations Act for the first time. Workers at partnerships, family businesses and sole traderships; outworkers; domestic workers; participants in the cash economy; and many others, could potentially be regulated by the Commonwealth, in some cases for the first time. Such breadth would rely, of course, on a liberal understanding of trade and commerce and what sufficiently affects it, but it could well herald a major increase in coverage, and an associated development in the way we think of labour law in Australia. Those that have in the past fallen through the gaps caused by the exigencies of the industrial power, and those that would continue to be systemically excluded under a corporations based-scheme, could for the first time find their place in the schema of employment rights and relationships.

5. Conclusion

I have argued that the costs, gaps and overlaps that characterise the Australian industrial relations system are primarily caused by an over-reliance upon the limited industrial power. The requirement that disputes be interstate results in a dual system of state and federal industrial jurisdiction; ‘conciliation and arbitration’ requires a particular form of dispute resolution; and ‘dispute’ prevents

\textsuperscript{228} See \textit{WR Act} s5.

\textsuperscript{229} See \textit{WR Act} s170CB(1)(d).

\textsuperscript{230} See \textit{WR Act} s5AA(3) and Pt VIB (certification by AIRC); s170VC(d)–(f) and Pt VID (approval by Employment Advocate).

\textsuperscript{231} See \textit{R v Wright; Ex parte Waterside Workers’ Federation of Australia} (1955) 93 CLR 528 at 539, 540 and 544 (Dixon CJ, McTiernan, Williams, Webb, Fullagar, Kitto & Taylor JJ); \textit{Seaman’s Union v Utah Development Co,} above n51 at 137–139 (Gibbs J) & 152, 154 (Mason J). Although this is probably the correct result, it does illustrate something important about our conception of labour, I think, that judges so quickly conclude that labour, as a ‘resource’, is a part of a commercial process (as opposed to a part of society).

\textsuperscript{232} See s170VA (‘constitutional trade and commerce’). Compare s127C(1)(b) (constitutional corporation) and discussion in \textit{Re Dingjan,} above n83.
the Parliament from directly and clearly setting out the rights of Australians as workers and employers. Though the *Workplace Relations Act* now relies upon a web of alternative constitutional provisions to ameliorate these problems, neither the corporations nor the external affairs power provide the Commonwealth with both broad coverage and policy flexibility.

Thus, the Commonwealth and the High Court should look afresh at industrial regulation under the trade and commerce power. Section 51(i) is a sleeping giant, left unconsidered for too long in times of rapid economic and industrial change. Most commentators now agree that the present construction is unsustainable and unsound. The ‘artifices of legal formalism’ that permit the narrow view must now be rejected. A flexible, broad, unitary industrial system is in sight.

---

233 Mason, ‘Minimalism, Monarchism or Metamorphism’, above n151 at 11.
Being Ms B: B, Autonomy and the Nature of Legal Regulation
DEREK MORGAN & KENNETH VEITCH*

Abstract
In this article, we question the apparent simplicity of medical law’s construction of ‘life and death’ cases as a clash between the sanctity of life principle and patient autonomy. Our main purpose in doing so is to try to understand more fully the nature of law’s regulation of the existence and non-existence of life. Specifically, we argue that, by broadening the understanding of autonomy in this area beyond a simple concern for patients’ rights and self-determination, to include a focus on the individual generally, it becomes possible to identify some of the legal practices that are central to the manner in which law regulates the threshold between life and death. Through an analysis of a recent case in English law — Re B (an adult: refusal of medical treatment) — (although Australian jurisdictions presently disclose no similar, authoritative case, ours presently is almost an arbitrary choice) — we demonstrate the central role played in this regulation by tests for mental capacity, questions of character, explanation, and imagination. We conclude that medical law, at least in this context, can be theorised as a normalising practice — one in which the determination of norms often occurs through patients.

* Cardiff Law School. An earlier version of this paper benefited from the reviews of two anonymous referees to whom we are indebted for critical and constructive comments.
1. **Introduction**

It is something of a puzzlement that Australia has no *definitive* case on death. Most common law jurisdictions have explored, in their superior courts in the past ten years, arguments that have engaged sanctity of life arguments and right to die claims. *Airedale National Health Service Trust v Bland*¹ and *Pretty v United Kingdom*² in England & Wales; *In Re Quinlan*,³ *Cruzan v Director, Missouri Department of Health*,⁴ *Washington v Glucksberg*⁵ and *Quill v Vacco*⁶ in the United States; *Auckland Health Board v JMB*⁷ in New Zealand; and *Rodriguez v British Columbia (Attorney General)*⁸ and *Nancy B v Hotel-Dieu de Quebec* in Canada⁹ have each engaged fundamental questions of death and dying. Australian jurisdictions have, of course, generated statutory law on assistance in dying and advance refusals of treatment,¹⁰ and the scholarly literature is large, and, in the case of Margaret Otlowski’s magnum opus¹¹ and Roger Magnusson’s ground-breaking study,¹² definitive. But how Australian courts would approach sanctity of life and right to die arguments remains moot. There would, however, be little doubt that the keen legal embrace of autonomy that has been characteristic of judicial involvement in matters of medical practice and death in recent years, would feature squarely.

It is that core ethical concept and its reception in the common law that we intend to explore in this paper. It is not, however, our intention to describe the different notions of autonomy that might be discerned in medical law, with a view to adopting, and arguing for, an endorsement of one definition rather than another. To date, the discussion of autonomy in the academic legal literature has predominantly been confined to the criticism of cases where the judiciary has not recognised or upheld the appeal to the autonomy, or rights, of patients. The subsequent arguments have called for the acknowledgment of more patient autonomy by the courts in future. In other words, the overriding concern has been a quantitative one — more autonomy; more patients’ rights; less professional medical power; more intervention by the courts in setting standards. In short, such arguments have, to a large degree, concerned the emancipation, through law, of patients from clinical capture by the medical model.

---

1 [1993] 1 All ER 821 (hereinafter *Bland*).
6 117 S Ct 2293.
7 [1993] 1 NZLR 235.
8 (1993) 82 BCLR (2d) 273 (SCC).
9 86 DLR (4th) 385.
10 Rights of the Terminally Ill Acts 1995, 1996 (NT); Medical Treatment Act 1988 (Vic) ss3–9; Natural Death Act 1988 (NT) ss3–7; Medical Treatment Act 1994 (ACT) ss4,5, 10–12, 20–23; Consent to Medical Treatment and Palliative Care Act 1995 (SA) ss7–11.
There is no doubt that these arguments have been important, and may, indeed, have contributed to the current judicial fixation with an idea of autonomy based on rights to self-determination. They have, however, tended to proceed in an either/or manner — *either* sanctity of life holds the winning hand *or* the individual’s right to self-determination ought to trump this traditional ethic. Here, we propose to investigate how the idea of autonomy lives within the common law when it comes to address the various problems created by medical practice and technology. To this end, our approach to autonomy is more concerned with the ethos that has developed recently within the common law that seeks to place more emphasis on the individual who comes into contact with medical practice. We are interested in autonomy as an instance of law’s turn towards the individual generally, and not solely, or even principally, with calls for patients’ rights and emancipation. As such, we intend to examine the legal practices that have grown up around autonomy and that help to determine the manner in which it operates in relation to the potential death of sick individuals. What do those practices impact upon and live off? We will seek to demonstrate that they exist in a *symbiotic* relationship with other factors, such as questions of identity and character, both of which are central to an understanding of the nature of legal regulation in cases involving life and death. If there is deeper resonance to the rhetorical chorus of ‘rights to die’, the same can be said of the rhetoric with which autonomy has come to occupy a central place in medical law. This essay is a preliminary examination of some of what informs this rhetoric, and we focus on the English case of *Re B (an adult: refusal of medical treatment)* to illustrate our argument.13

2. **Death, Dying, Rights**

‘Right to die’ arguments are, on one account, complex, confused, confusing, and controversial. Leon Kass has referred to what he believes is the lexical hopelessness of a narrowly conceived, literally constructed notion of a right to die.14 He has written that: ‘I do not think that the language and approach of rights are well suited either to sound personal decision-making or to sensible public policy in this very difficult and troubling matter.’15 In arguing from this premise, he concludes that there is no firm philosophical or legal argument for a ‘right to die’: ‘My body and my life, while mine to use, are not mine to dispose of.’16 And John Finnis has dismissed as mere sloganising the use of a term such as ‘euthanasia’ which, he claims, is devised ‘for service in a rhetoric of persuasion’ because it ‘has no generally accepted and philosophically warranted core of meaning.’17

Max Charlesworth, on the other hand, has identified much of the prose for which the ‘right to die’ stands as a shorthand expression. Seen as the expanded notion of controlling the manner and the means, the geography and the grail, of

---

13 [2002] 2 All ER 449 (hereinafter *Re B*); Dame Elizabeth Butler-Sloss P handed down judgment on 22 March 2002.
15 Id at 34.
16 Id at 39.
one’s death, ‘this developing recognition of the right of a person freely to
determine and control, so far as is possible, the mode of his death’18 is part of the
attempt to recapture the right to preside at one’s death, the loss of which is
bemoaned by writers as diverse as Ivan Illich and John Gray. Illich, for example,
has complained that: ‘[S]ociety, acting through the medical system, decides when
and after what indignities and mutilations [the sick person] shall die … Western
man has lost the right to preside at his act of dying … Mechanical death has
conquered and destroyed all other deaths.’19

Concern with death, dying and euthanasia is really nothing new; but has there
been a change in the engagement?20 There are two central points to the movement
of the compass which are definitive — one as a response to movements in the
other. There has been, first, as remarked upon by so many before, the
medicalisation of death (if not life more generally). The medical management of
death and dying is perhaps one of the most salient changes in the general practice
of Western medicine in the past century. More and more people now die after an
explicit decision has been made, either to withdraw or not to start treatment. Much
modern medicine ‘is pathological in its denial of death’ and reflects the broader
culture of which it is a part ‘in refusing to recognise that we may thrive in dying,
even as our souls may perish in senseless longevity.’21 Fear of dying, of the
possible manner of death and indeed of death itself, are important parts of the
human condition.22 Paul Ramsay suggests that awareness of dying constitutes an
experience of ultimate indignity in, and to, the self who is dying.23 In our dread we
are capable of doing much harm; harm that may extend to patients, their families,
the medical team and society at large.

17 John Finnis, ‘A Philosophical Case Against Euthanasia’ in John Keown (ed), Euthanasia
Examined: Ethical, Clinical and Legal Perspectives (1995) 23–35 at 23. Or, as otherwise put,
‘When I use a word,’ Humpty Dumpty said, in a rather scornful tone, ‘it means just what I
choose it to mean – neither more nor less.’ ‘The question is,’ said Alice, ‘whether you can make
words mean so many different things.’ ‘The question is,’ said Humpty Dumpty, ‘which is to be
master – that’s all.’ ‘Humpty Dumpty’ in Lewis Carroll, Through the Looking Glass printed in
Martin Gardner (ed), The Annotated Alice (1965) at 268–69.
See Frederick Nietzsche, ‘Twilight of the Idols’ in The Complete Works of Frederick Nietzsche
at 88, quoted in Ronald Dworkin, Life’s Dominion: An Argument about Abortion and
Euthanasia (1993) at 212.
20 Daniel Callaghan offers interesting reflections on this question, while concluding that ‘[h]ere
is no clear and obvious explanation…’. See his ‘Foreword’ in Keown, above n4 at xiv.
at 167.
22 John Saunders, ‘Medical Futility: CPR’ in Robert Lee & Derek Morgan (eds), Death Rites: Law
and Ethics at the End of Life (1994) at 72–90. The Royal College of Nursing in evidence to the
House of Lords Select Committee on Medical Ethics suggested that ‘… many people are not
necessarily afraid of death, but are afraid of the manner of death.’ House of Lords Select
Committee Report of the House of Lords Select Committee on Medical Ethics, HL Papers 21–I
(1994) at [187].
Secondly, this medicalisation of death and dying has been one factor that has created the conditions for the emergence of a ‘right to die’ discourse. And, while it may be clear that such claims do not enjoy any extensive philosophically agreed core of meaning, we all know what they mean, and what others mean when they deploy them. The rhetoric of rights to die, like all advertising slogans and expressions, both debases as well as encapsulates something larger. It is syntactical and symbolic shorthand about timing and temperament, attendants and attitude, place and purposes.

The House of Lords in its deliberative capacity has set itself against euthanasia, and in its judicial capacity against relaxation of the rule prohibiting assisted suicide. Lord Steyn in Ex parte Pretty summarised the present position in English law:

By virtue of legislation suicide is no longer an offence … Mercy killing in the form of euthanasia is murder and assisted suicide is a statutory offence … A competent patient cannot be compelled to undergo life saving treatment. Under the double effect principle medical treatment may be administered to a terminally ill patient to alleviate pain although it may hasten death. This principle entails a distinction between foreseeing an outcome and intending it. [In Bland] the House of Lords held that under judicial control it was permissible to cease to take active steps to keep a person in a permanent vegetative state alive. It involved the notion of a distinction between doctors killing a patient and letting him die. These are at present the only inroads on the sanctity of life principle in English law.

Bland recognised, controversially, that the sanctity of life in English law was not an absolute value. Lord Goff said that ‘the sanctity of life must yield to the principle of self determination,’ and Lord Keith that ‘a person is completely at liberty to decline to undergo treatment even if the result of his doing so is that he will die.’ Ex parte Pretty; as Michael Freeman and Richard Tur have argued, is rapidly coming to seem to be the anomaly in English law. To that extent the Chinese walls that have been constructed around these different forms of dying and bringing about death — entailing a distinction between foreseeing an outcome and intending it — seem to be closing in. Both death and the ‘right’ to it would appear to be finding legal footholds as the courts are increasingly forced to reveal their hand. Yet, in English law, the discourse of the ‘right to die’ has largely been absent as a mechanism to be used to justify the progressive erosion of the traditional value.

---

24 See Charlesworth, above n5 at 34.
25 House of Lords Select Committee, above n9.
26 R v DPP; Ex parte Pretty [2001] UKHL 61 (hereinafter Ex parte Pretty).
27 [2001] UKHL 61 at [55].
29 Id at 857.
placed on life. It is, rather, through the idea of autonomy that substantive aspects of ‘right to die’ claims have featured in the common law. That, again, is the focus of this examination.

3. **Chronicle of a Death Unfolding — On the Woman B**

We glean that the case of Ms B is no ordinary case from the way in which Butler-Sloss P frames the facts; the case involves ‘the tragic story of an able and talented woman of 43 who has suffered a devastating illness which has rendered her a tetraplegic. Her expressed wish, both verbally and in a written ‘advance directive,’ is not to be kept artificially alive by the use of a ventilator.’ Neurological surgery was successful to the extent that it allowed her to move her head and speak, but thereafter, she made several requests for ventilation to be withdrawn, and, as a result, repeated assessments were made of her mental capacity to decide to refuse the treatment. Eventually, although it was decided that she did have this capacity, (and the hospital treated her as such), the clinicians were not prepared to withdraw the ventilator. The doctors’ dilemma — the court describes their position in robust terms as having been unjust; the Trust, on the other hand, failed in its responsibilities to them as clinicians — is summed up in Butler-Sloss P’s review of the evidence of the lead clinician in the case:

>[Dr C] had studied and spent her professional life trying to do her best to improve and preserve life. She did not feel able to agree with simply switching off Ms B’s ventilation. She would not have been able to do it. She felt she was being asked to kill Ms B.]

In her application, B had sought a declaration that she indeed had mental capacity to refuse treatment and she asked the court to declare that her continued artificial ventilation amounted to an unlawful trespass to her person. Although popularly phrased a case involving a ‘right to die’, Re B apparently involved the legally less prosaic but nonetheless potent choice to accept or refuse medical treatment. It is not a case about euthanasia or physician assisted suicide, although there may have been times when Ms B might have been forgiven for wishing that it was. This is why it is little short of idle to proclaim that Re B is a case merely of refusal of treatment; of course it is. But it is not so much a question about what B means — in one way that is quite a straightforward enquiry — but rather where it fits. It would be more than idle to pretend that the case does not arise within and against

---

32 On living wills, or advance statements, see British Medical Association, *Statement About Advance Directives* (1992); *Re C* [1994] 1 WLR 290; *Bland*, above n1 at 860a–b (Lord Keith) and 866b–e (Lord Goff). For the Court of Appeal’s consideration of this point, see [1993] 1 All ER 821 at 843a (Lady Justice Butler–Sloss), 835–36 (Sir Thomas Bingham) and 852–54 (Lord Justice Hoffman). See also *Re AK* (2000) 58 BMRL 151.
33 *Re B*, above n13 at 452.
34 Id at 463.
35 Of course, there is a limitation to the relief here; the declaration can, by definition, only address her mental capacity as at the date that the declaration is granted. It cannot, for example, declare what her future state might be, or if assessed, might be assessed to be.
the background of what Roger Magnusson has called, in a related context, the erosion — the ‘slow death’ — of the doctrine of the sanctity of life.36

The public reportage of the case would make it appear quite straightforwardly as an argument between Ms B — who argued that the law provided that her immediate death should be a foregone conclusion — and her doctors, cast in the role as engineers of the human soul, who argued that that should be a conclusion foregone. A careful reading of Butler-Sloss P’s decision, however, makes it clear that not only were the factual issues involved more complex than this but also that the legal sequelae that flow from this complexity call for a more nuanced understanding. Her judgment suggests that a doctor’s legal duties flow differently for a patient deemed to be mentally ‘competent’ and one who is assessed as mentally ‘incompetent’, and that the complexity is deepened by the possibility of bringing both the question of a patient’s ambivalence about the withdrawal of medical treatment, and the failure to experience alternative forms of treatment, to bear on such assessments. Butler-Sloss P clarifies that the doctors must, legally, make several inter-related judgments and that, on the basis of those judgments, depends the nature of their legal duties. Far, then, from being a simple monotone case, Re B involves not just one but many shades of grey against which background even the most powerful argument might be occluded or dimmed. And, while the case would seem to be a further entry in the library recording the clash between traditional Hippocratic values and modern medical law’s apparent insistence on, and regard for, autonomy and rights, we will argue that there is much more going on beneath the surface that demands a re-orientation of our standard ways of thinking.

4. Autonomy and Mental Capacity

Lord Justice Ward in Re A (Children)37 averred that ‘deciding disputed matters of life and death is surely and pre-eminently a matter for a court of law to judge.’38 It followed that:

This court is a court of law, not of morals, and our task has been to find, and our duty is then to apply the relevant principles of law to the situation before us — a situation which is quite unique.39

Many lawyers would unhesitatingly endorse such powerful sentiments. We want to suggest, however, that they might also engender certain unease. That doubt might be prompted by asking what characteristics a court of law possesses that ranks it above all other means a society could offer for resolving a dispute concerning life and death.40 And, of course, the answer traditionally afforded might well be that it is the impartiality of law — to be guaranteed through the neutrality of legal principles — that marks out courts as the privileged social site

36 See Magnusson, above, n12 at 54 et seq.
37 [2000] 4 All ER 961.
38 Id at 968.
39 Id at 969.
for ruling on any such disputes. The consequence of this self-proclaimed social
pre-eminence is that, while life and death may engage a range of values and attract
a plethora of discussions — moral, philosophical, historical, sociological, political,
scientific, to name but a few — disputes about them are best thought of in legal
terms.

This judicial need easily and clearly to distinguish between, to separate, law and
morals, encounters and embraces some difficulty in medical law, given the
historical foundations of medicine as an ethical practice.41 Ian Kennedy, amongst
others, has reminded us of the moral basis of medical practice and the extent to
which this necessarily involves its practitioners in the world of practical, normative
ethics for which their education and training has traditionally little prepared them.
Although this has not gone without drawing stinging ripostes42 and more
thoughtful, reasoned responses,43 we have — properly — insisted on the practice
of modern medicine involving, implicating, indeed being an ethical one. Thus, it
can really come as no surprise (indeed it may be thought to be a cause of
celebration) when debates about contested practices within the art of medicine
come to be conducted in essentially moral terms, and that changes in medical
practice come to be rehearsed in moral language. Are debates in law, then, so
distinct? Not all members of the judiciary are so concerned to fence off the
common law from all possible incursions by moral reasoning when it is brought to
judgment in contested areas or instances of medical practice.44 Indeed, Lord
Justice Ward’s own essay in utilitarian ethics, in which his judgment in Re A
(Children) is embedded, demonstrates this only too well. It is important to stress,
however, that, in medical law, this play between morals and law is not confined to
grand statements that seek to preserve the purity of the common law. Rather, the
implicit characteristic of maintaining a distinction between the technical (which
seeks to ensure impartiality) and the moral (associated with complexity, difference,
and values) has filtered down from such elevated levels to occupy a more
significant position in individual cases. One such example, the legal distinction
made between autonomy and mental capacity, can be found in Re B. To make any
proper sense of this distinction, however, we need briefly to survey its history.

The idea of autonomy in medical law has traditionally manifested itself as an
end-point; it is a goal, the attainment of which is becoming ever more important.
In 1992, Lord Donaldson MR delivered a significant judgment that contained a re-
statement45 of the underlying value of autonomy and a (slight) amplification of its
nature:

---

40 A question addressed, for example, in Guido Calabresi’s classic polemic with Phillip Bobbitt,
41 On this, see, for example, Albert Jonsen, A Short History of Medical Ethics (2000).
43 Of which the most accessible is Anne Maclean The Elimination of Morality (1993) at 187.
44 See, for example, the judgment of Lord Steyn in McFarlane and Another v Tayside Health
Board [1999] 4 All ER 961.
45 The importance of autonomy as a value in medical law extends back at least to Cardozo J’s
judgment in Schloendoff v Society of New York Hospital 105 NE 92 (NY, 1914).
An adult patient...who suffers from no mental incapacity has an absolute right to choose whether to consent to medical treatment, to refuse it or to choose one rather than another of the treatments being offered.... This right of choice is not limited to decisions which others might regard as sensible. It exists notwithstanding that the reasons for making the choice are rational, irrational, unknown or even non-existent.46

His comments appear to bolster the esteem in which the courts hold the sovereignty of the patient. As Martyn Evans has summarised it: ‘When doctors acknowledge a patient’s freedom to make “bad” choices seemingly against his or her clinical interests, they acknowledge the triumphing of science by ethics and they confront the uneasy blend of the biological and the biographical in the human person.’47 Donaldson’s reference to the option of giving reasons for decisions purports to place the individual with ‘no mental incapacity’ beyond the scope of inquiry. The problem, however, has been this — how do we arrive at this ‘good’? By what method(s) are we able to turn to an individual and say: ‘You can now exercise your autonomy’? If the end-point in law has its roots in ethics, the journey towards it can, by contrast, be seen to pride itself on the apparently neutral technicalities of procedure. It has been this urgency to create a noticeable division between the nature of the end-point and the method of arriving at it, that has characterised the common law’s approach to autonomy in the health care context.

The assessment of individuals lies at the heart of the journey to medical law’s end-point of autonomy. This assessment manifests itself in various forms, the most common of which is the extent to which individuals can be said to have ‘mental capacity’ or to be ‘mentally competent’. While the ideas of capacity and competency have common features (such as ability and talent), it is usually stressed that their use in the legal context refers to a particular notion of competency or capacity. In the medico-legal context, for an adult to be legally competent, the main factor to be considered is the state of his or her mind.48 The answer to the question: ‘Should this individual be allowed to exercise his or her right to self-determination?’, is discovered by asking, initially: ‘To what extent

46 Re T (Adult: Refusal of Treatment) [1993] Fam 95 at 102. The latter comment in relation to reasons had first been noted by Lord Templeman in Sidaway v Board of Governors of the Bethlem Royal Hospital and the Maudsley Hospital and Others [1985] AC 871 at 904: ‘If the doctor making a balanced judgment advises the patient to submit to the operation, the patient is entitled to reject that advice for reasons which are rational, or irrational, or for no reason.’ Of course, there may be both logical and morally significant reasons – not all related to autonomy – for distinguishing between making a choice and reasons for rejecting advice.

47 Martyn Evans, ‘Philosophy and the Medical Humanities’ in Martyn Evans & Ilora Finlay (eds), Medical Humanities (2001) at 251.

48 This is the approach commonly adopted in Australian statutory tests of understanding. For example, the Powers of Attorney Act 1998 (Qld) sched 3 and the Guardianship and Administration Act 2000 (Qld) sched 4 define capacity as ‘(a) understanding the nature and effect of decisions about the matter; and (b) freely and voluntarily making decisions about the matter; and (c) communicating the decisions in some way.’ Similar definitions can also be found in the Guardianship Act 1987 (NSW), s33(2); Guardianship and Administration Act 1995 (Tas) s36(2) and Guardianship and Administration Act 1986 (Vic), s36(2). See also Re Bridges [2001] 1 Qd R 574.
does he or she satisfy the requirements of the decision-making process set out by
the common law?’. This process takes the form of a series of three questions, the
specific purpose of which is to decide whether the individual has understood the
nature, purpose and effects of a proposed medical treatment. If he or she has done
so, mental capacity is deemed to exist and nothing ought to stand in the way of the
individual’s exercise of autonomy. The three elements to be satisfied are the
following: (a) has the individual been able to comprehend and retain information
provided by the physician regarding medical treatment?; (b) has he or she believed
that information?; and, (c) has he or she weighed the information, balancing risks
and needs, when arriving at his or her choice?49

In Re B, this legal distinction between autonomy and mental capacity was, once
again, central to the management of the dispute. From the outset, Butler-Sloss P
was careful to stress that this was a case about ascertaining the presence or absence
of an individual’s mental capacity to make decisions about medical treatment.
This, she said, was not to be confused with the nature of the patient’s decision,
however grave the consequences of Ms B’s refusal of treatment would be. The
patient’s decision may ‘reflect a difference in values rather than an absence of
competence.’50 In other words, doctors, and presumably the courts, must exclude
consideration of what a patient has decided when trying to establish whether he or
she has sufficient mental capacity to make a decision about treatment.

There are several problems with this legal attempt to create a distinction
between mental capacity and autonomy. Those problems strike not only at the root
of one of the guiding distinctions in medical law, but, more importantly, demand a
re-orientation in our thinking about the nature and operation of law in this area.

A. Capacity and Decisions

The first difficulty with the distinction arises even before beginning to consider the
elements of the test for mental capacity or, to put it more realistically, the
assessment of individuals. The problem is this: how can you ascertain a patient’s
mental capacity without being concerned with the nature of what he or she
decides? The general point is captured by Ian Kennedy:

When devising tests for incapacity, are we talking about the patient or are we
talking about a decision taken by the patient? The answer must be that we are
concerned with both. We are concerned to establish that the patient meets certain
criteria and one of the ways in which we seek to determine this is by examining
the decision reached by the patient. The two, the patient and the decision, are
inextricably intertwined. The trouble is that the moment we admit this, that the
content of a patient’s decision is relevant in the determination of capacity, we face
the problem of autonomy simply being overwhelmed by paternalism .... It is its

49 See Re C (Adult: Refusal of Medical Treatment) [1994] 1 All ER 819 (Thorpe J). It should be
noted that, while this test to ascertain mental capacity was endorsed by Butler-Sloss LJ in Re
MB (An Adult: Medical Treatment) [1997] 2 FCR 541, her manner of formulating it seems to
have rendered the second stage inapplicable.

50 Re B, above n13 at 450.
desire to avoid this (inevitable) conflation of decision with decision-maker that led the Court of Appeal [in Re MB]51 to want to be seen to nail its colours to the mast of patient autonomy.52

While Kennedy sees the patient’s decision as one possible means by which to examine mental capacity, we would suggest that it is the very nature of the patient’s decision that results in the need to question and assess his or her mental capacity to make decisions. What would be the point of tests to establish mental capacity if the nature of one’s decisions was beyond investigation? If Ms B had agreed with the medical staff to continue with ventilation, there would have been no need to question, or assess, her capacity. So, in fact, contrary to what Butler-Sloss P argues, and although supported by a line of judicial authority, the first pre-requisite for the establishment of tests for mental capacity is revealed as a consideration of the nature of a patient’s decision. In other words, a decision is required, and one whose nature offends medical practice, the ethics of the medical profession, the personal values of the medical staff, or whatever. It is important, then, to note that tests for mental capacity arise as a result of the nature of the patient’s decision.

This conflation of the nature of decisions and the assessment of mental capacity can be seen further in the following statement from Butler-Sloss P’s judgment in Re B:

I shall … have to consider in some detail her ability to make decisions and in particular the fundamental decision whether to require the removal of the artificial ventilation keeping her alive. It is important to underline that I am not asked directly to decide whether Ms B lives or dies but whether she, herself, is legally competent to make that decision. [Emphasis added.]53

The question to be asked is this: is this woman legally competent to decide whether she lives or dies? In other words, the assessment of her mental capacity is to be measured against the ‘fundamental’ nature of the decision to be made — that between life and death. But the point is, surely, that she has already made such a decision. She clearly, then, has the ability to make the decision. What the law is interested in, however, is something completely different — that is, does she have the ability to decide as she has. This ability has nothing to do with actually making a decision. If it had, there would have been little reason to seek the intervention of the courts. Rather, it is, fundamentally, to do with ruling on the existence or destruction of life as such and, specifically, whether Ms B can persuade a judge that she deserves to flout the importance that law often places on the maintenance of human life. Thus, the real point of legal tests for mental capacity seems not to be to assess some projected future or, indeed, past ability to make a choice between life and death, but to assess whether the person making that decision can construct a convincing case why he or she reaches the standard of the ‘ability’ that law expects in such circumstances. The nature of his or her decision is crucial in that it triggers this type of assessment.

51 Re MB, above n49.
53 Re B, above n13 at 454–455.
B. Accounting for Reasons

The second difficulty with the legal distinction between tests for mental capacity and the nature of patients’ decisions (or, the ‘autonomy’ stage) can be seen by returning to Lord Donaldson’s statement above in Re T. He stresses that adult patients who are not assessed as lacking mental capacity may choose freely with regard to medical treatment that has been offered to them. There is no need for them to prove to others that their decisions are reasonable; indeed, there is no obligation to offer reasons for their decisions. But this ‘freedom’ and, in particular, the reference to irrationality and the non-existence of reasons, sits uneasily with the final criteria of the three-stage legal test for determining mental capacity. The problem revolves around the question why, if an individual need give no reasons for making a decision, he or she must demonstrate that they have taken account of anything when making that decision? In other words, there is no need to give reasons for a decision, but it is necessary to show that you have performed a reasoning process in arriving at a decision. The important point is that the last stage of the test for mental capacity involves deciding or choosing, and the method of arriving at that decision or choice. The legal procedure as it currently stands would seem to read something as follows: (a) decision (effectively refusal of medical treatment); (b) assessment of mental capacity, including the weighing of information provided and the balancing of risks and needs when arriving at a decision; (c) finding of no mental incapacity; (d) the individual has a right to choose or make a decision for whatever, or no, reason at all. Lord Donaldson’s frequently cited words about individuals being able to decide as they see fit for whatever or no reason at all, would appear to be superfluous given that decisions must, in the first instance, conform to a certain procedure, one of the purposes of which is to assess how patients make decisions — how they weigh information, balance it and so on. If this is so, then what interests those assessing mental capacity is, as we shall soon see, the whole decision-making, or reasoning, process, including the various factors (such as values and beliefs) against which the treatment information is weighed, balanced, and so on. The procedure of assessment cannot conveniently omit consideration of those wider aspects.

This notion of autonomy as it currently exists in medical law cannot, then, be thought of as a separate stage that follows on from assessments of the mental capacity of patients. Rather, the extent to which the former is deemed to exist at all is wholly dependent on the procedure adopted for assessing the mental capacity of individuals to make decisions they have already taken. Before going on to discuss the consequences of this conflation, it is worth noting one further important aspect of Lord Donaldson’s judgment in Re T.

As we have noted, the law envisages the assessment of individuals as occurring at the level of mental capacity. The nature of decisions is considered to be distinct and, assuming mental capacity is deemed to exist, beyond investigation. But Lord Donaldson’s judgment makes it clear that, even where the individual is considered not to lack mental capacity, and therefore has a right to make a decision about treatment for whatever, or no, reason, assessment of the way in which the individual chooses to exercise this right is still possible:
Just because adults have the right to choose, it does not follow that they have in fact exercised that right. Determining whether or not they have done so is a quite different and sometimes difficult matter. And if it is clear that they have exercised their right of choice, problems can still arise in determining what precisely they have chosen. [Emphasis added.]54

How does the common law seek to justify questioning the choices that it has declared individuals are capable of making?

It is important to note that Lord Donaldson takes care not to base his qualification of the right to choose on any criticism that could be made of the nature or merits of the patient’s reasons for choosing to consent to, or refuse, medical treatment — how could he when, by his own ‘test’, reasons for refusal need not be rational or even exist?55 Instead, as his judgment on the facts of this particular case reveals, it is to ostensibly technical matters, such as the extent of third party influence on the patient’s decision and the scope of a patient’s refusal, that we ought to turn to find reasons for overriding his or her right to choose. So, where patients have exercised their right to choose, and Lord Donaldson comments that ‘problems can still arise in determining what precisely they have chosen’, he is referring, in the instant context, to whether T’s refusal of a blood transfusion was intended to apply to a changed set of medical circumstances. Thus, what appears is a second level of inquiry — namely, an assessment of the legally autonomous individual’s manner of choosing. How did this individual choose? What did he or she mean to express in that choice? What were the boundaries of its application? Despite the existence of a right to choose, judicial discretion can still intervene to question the nature of its exercise in an individual case. What is interesting is to probe the origin of this discretion; for if a right of choice has been exercised by an individual, it would seem strange to ascribe, as Lord Donaldson tries to do, the need to question it to the indecisiveness of that individual.

In order to appreciate the beginnings of this doubt, it is necessary to emphasise the circumstances in which it comes to exist. We have now reached the end-point of autonomy, where the right of choice is viewed as an important ethical value. As such, Lord Donaldson acknowledges that the woman’s appeal is not only concerned with the right to choose to refuse medical treatment; it is also ‘about the “right to choose how to live.”’56 It is unsurprising, then, to find that Lord Donaldson hints at a broader conflict of interests:

The patient’s interest consists of his right to self-determination — his right to live his own life how he wishes, even if it will damage his health or lead to his premature death. Society’s interest is in upholding the concept that all human life is sacred and that it should be preserved if at all possible. It is well established that

54 Re T, above n46 at 102 (Lord Donaldson MR).
55 It is, however, illuminating to note a comment made by Lord Donaldson in another ‘competency’ case decided only one month before Re T – ‘I personally consider that religious or other beliefs which bar any medical treatment or treatment of particular kinds are irrational…’. See Re W (A Minor) (Medical Treatment: Court’s Jurisdiction) [1993] Fam 64 at 80.
56 Re T, above n46 at 102 (Lord Donaldson MR).
in the ultimate the right of the individual is paramount. But this merely shifts the
problem where the conflict occurs and calls for a very careful examination of
whether, and if so the way in which, the individual is exercising that right. In case
of doubt, that doubt falls to be resolved in favour of the preservation of life for if
the individual is to override the public interest, he must do so in clear terms.
[Emphasis added.]

This classic conflict between two public interests (the interest in rights to self-
determination and the interest in the preservation of life) is interesting because, if
the patient is declared as having a right to choose, then no such conflict ought to
exist at all. If, as Lord Donaldson argues, there is no need for reasons, far less
rational ones, to be given for refusing treatment, why should any further ‘careful
examination’ of the manner in which an individual decides be necessary? This
question is an interesting and difficult one that calls for much deeper examination.
For the present, we will observe only that the assessment of the way in which an
individual chooses to exercise his or her right to choose does not cease even after
it has been decided that he or she does, in fact, have such a right.

5. Explanation, Character and Imagination

This convergence of the technical (tests for mental capacity) and the evaluative
(the nature of decisions or the ‘autonomy’ stage) produces situations that have an
important bearing on understanding the nature of medical law. We will first outline
some of those situations and, in the following section, discuss how they are
symptomatic of the way in which law regulates questions of life and death.

As the assessment of mental capacity and the nature of patients’ decisions
merge, so individuals must explain the reasons for their decisions. For example, in
Re B — through questions designed to assess whether she had been ambivalent in
her views about the withdrawal of ventilation, and why she did not want the
proffered one-way weaning programme — the patient was required to explain
what she thought about dying and the manner in which this would occur. In order
to justify her choice, she had to speak not only of her decision and the reasons
behind it, but also of her suffering. This general requirement is captured by
Marinos Diamantides, discussing the case of State of Tennessee v Northern:
This, he claims, appears to have made the right to refuse treatment ‘subject to an
existential requirement: the patient must take a pro-active attitude towards his or
her suffering. They must speak “of it”…’

57 Id at 112.
58 The purpose of this programme is gradually to reduce the number of breaths supplied by the
ventilator, thereby allowing the patient’s body to become accustomed to breathing without
assistance again. If difficulty in breathing occurs, the patient is not given artificial ventilation
again, but, rather, is only sedated.
59 563 SW 2d 197 (1978) (Tenn Ct App).
at 64.
This is also illustrated in Re B, where it becomes impossible for Ms B to confine her explanations about the possible treatment alternatives to the level of technical benefit or risk. If the third vector of the legal test for mental capacity envisaged such a neutral assessment, for Ms B, to talk about ventilators and one-way weaning programmes was to talk about death, her values, and beliefs. So when, for example, Mr Francis QC, representing the hospital Trust, asked her whether what she wished for was to die or not to remain alive in her present condition, part of her explanation was that she found the idea of living in this condition intolerable. It was a question directed to uncovering the root of the nature of her decision.

This requirement for Ms B to speak of the nature of her decision led to an assessment of how well, in the judge’s opinion, she had spoken of it; moreover, Ms B was fortunate in remaining competent to speak to the court. This also required a judgment of her character in deciding whether she was to be considered as having mental capacity.61 This is illustrated by Butler-Sloss P’s comments in her judgment:

Her wishes were clear and well-expressed. She had clearly done a considerable amount of investigation and was extremely well-informed about her condition. She has retained a sense of humour and, despite her feelings of frustration and irritation which she expressed in her oral evidence, a considerable degree of insight into the problems caused to the hospital clinicians and nursing staff by her decision not to remain on artificial ventilation. She is, in my judgment, an exceptionally impressive witness. Subject to the crucial evidence of the consultant psychiatrists, she appears to me to demonstrate a very high standard of mental competence, intelligence and ability.62

Who is this woman? Does she communicate and, if she does, does she do so effectively? Does she incorporate within her communication more than simply references to her own predicament? From Butler-Sloss P’s summary, it is clear that Ms B is considered not only to be articulate, but also knowledgeable about her condition, capable of responding to her situation with humour, and appreciative of the problems her decision had caused to those caring for her. This latter aspect seemed particularly important to the judge. It was something that needed to be overcome or spoken of in properly lucid terms. Its negative implication had to be sufficiently compensated for by her general attitude and demeanour, something that she clearly managed to do owing, amongst other things, to her insight, intelligence, and ability. The overall impression given by the judge was that Ms B had reached certain unspecified standards of character and persuasiveness that

61 Assessments of individuals’ characters and personalities have also been important in other areas of medical law, particularly where allegations of negligence are made against doctors who have not disclosed information about the potential risks of medical treatment. In this regard, see, for example, Sidaway v Board of Governors of the Bethlem Royal Hospital and the Maudsley Hospital and Others [1985] AC 871 and McAllister v Lewisham and North Southwark Health Authority and Others [1994] 5 Med LR 343.
62 Re B, above n13 at 462.
made her deserving of having her decision upheld. What is clear is that, if such a grave decision were to be condoned, there needed to be much more than a bare explanation of the risks and benefits of medical treatment. In other words, here, the law, in ascertaining the existence of mental capacity, demands much more of the patient than its own tests appear to require.

But it was not merely the court’s assessment of this woman that mattered. Given that such cases required ‘the highest degree of scrutiny’, the evidence of the two consultant psychiatrists called to testify was, according to Butler-Sloss P, ‘crucial’. While there was much emphasis placed on the need to respect the values individuals often express through their decisions about medical treatment, and to reiterate that questions of mental capacity were to be distinguished from the nature of individuals’ decisions, the importance attached to the evidence of psychiatrists in the ascertainment of mental capacity confirms the depth and range of possible factors the law considers may validly impinge on the state of an individual’s mind. Questions were raised, for example, regarding the possibility of psychological regression, the patient’s childhood experiences, and whether she suffered from mental illness. Thus, while those temporary factors were not sufficiently evident in this case to render B mentally incompetent to refuse medical treatment, the broader significance lies in their validity to the legal process, and, consequently, the importance of the expert to this process.

The President’s stress on the division between assessing people’s mental capacity and the way in which they choose their own ends was, in part, intended to define the proper role of the medical expert in such matters. The legitimate authority of the expert is to be confined to the technical and neutral matter of assessing the individual’s mind with a view to ruling on capacity. However, the psychiatrist is not only involved in assessing the extent to which individuals can understand and retain specific treatment information, believe it, and weigh it in arriving at a choice; he or she is also concerned to assess the possible causes of why individuals decide as they do. Is there something in this person’s past (an incident or experience), for example, that might explain their decision? But to ask such a question with the purpose of establishing the presence or absence of mental capacity is to find oneself in the role of examining lives, the courses they have taken, the influences that have shaped them, the decisions that have affected them. In other words, the investigation of such factors has, as its purpose, the identification of any untoward events in the past that may explain why the individual chooses as he or she does. It is the nature of Ms B’s decision — choosing death over life — that is, once more, central to the need to involve psychiatrists in the first place, and the ‘crucial’ nature of psychiatric evidence can surely only rest in the extent to which it can uncover reasons or causes in the course of an individual’s life that may help to explain such an odd decision. Once again, the law’s reliance on this type of inquiry obviously enlarges the scope of the factors that can be taken into account in assessing mental capacity and tends to confirm the suspicion that there is much more at stake here than the assessment of Ms B’s ability to weigh the risks and benefits of medical treatment.

63 Id at 468.
A further result of the conflation of mental capacity and autonomy is the emergence of the importance of imagination — something that played a significant part in Butler-Sloss P’s judgment. One of the expert psychiatrists, Dr Sensky, reiterated the importance of distinguishing the patient’s decision from the assessment of her mental capacity. It was imperative, he said, to begin by focusing on the individual and his or her capacity to make decisions. In so far as values were concerned, the patient’s subjective values had to be taken into account and respected. In order to clarify his point, Dr Sensky referred the court to Kim Atkins’ paper ‘Autonomy and the Subjective Character of Experience’, on which Butler-Sloss P was later to rely.

The thrust of Atkins’ paper — itself based on Thomas Nagel’s seminal essay ‘What is it like to be a bat?’ — is that, while human beings experience and perceive in a similar way to one another, there is always something irreducibly unique about their experiences and perceptions. Drawing on Nagel, Atkins argues that it is this uniqueness that grounds both the subjective character of experience and the value of respecting the autonomy of individuals. Given the nature of this experience, it is impossible to know objectively what it would be like to be another human being. We can, however, imagine not just what it would be like for me to be in the position of another human being in a specific set of circumstances, but, also, what it would be like to be that other human being in that same set of circumstances. Atkins argues that the importance of such imagination stems from the need ‘to act from respect for that [other] person’s autonomy’.

The hypothesis, Atkins’s hypothesis that Butler-Sloss P endorses, as we might indeed have imagined, has triggered an enormous literature in the philosophy of mind. It invokes the problem of ‘other minds’ — the philosophical proposition, a very Cartesian notion, that the mental contents (what others think) is a mystery to us and therefore ‘internal’. This view suggests that others always remain a mystery to us and that we can only hazard, or guess, what is in their mind. Butler-Sloss P’s response to this conundrum is to say that we can resolve this problem with empathy. This gets us some way, we believe, to an understanding of one of the ways in which this problem is falsely set up. We do, in fact, operate in a world of common meanings, of linguistic signs, of cultural propositions, which means that we do have access to the embodied and emotional nature of reason, even of another; reading people, then, is inevitable, albeit here, unnecessary.

This thesis, however, adds nothing to the process of resolving Ms B’s case. It is not obvious that the argument is being used or deployed in a way that Atkins would have intended for it. This is clear from both the example she uses in discussion — a patient on a right ventricular assist device ‘where the ill person is

65 Thomas Nagel, ‘What is it Like to be a Bat?’ in Thomas Nagel (ed), Mortal Questions (1979) at 165-180.
66 Atkins, above n64 at 78.
67 Ngaire Naffine, of the University of Adelaide, alerted us to this.
68 See Mary Warnock, Imagination and Time (1994) for an accessible discussion.
incapable of exercising his or her autonomy’ [Emphasis added.]⁶⁹ — and in her conclusion to the paper:

There is, for each of us, something that it is like to be “me”, and it is called the subjective character of experience. Our subjectivity grounds the value we attach to persons and personal autonomy. When we are faced with the regrettable position of having to make dire decisions on behalf of another, the only way to act so as to respect that person’s autonomy is to promote a consideration of that person’s subjective perspective. [Emphasis added.]⁷⁰

Atkins, therefore, envisages the thesis of the subjective character of experience as applying to situations, unlike that of Ms B, where individuals are clearly unable to communicate their views directly (for example, where they are unconscious), either contemporaneously or through the medium of an applicable advance statement. And here, as we have shown above, the legal conditions for believing that the individual’s very choice may be invoked are likely to be absent. ‘Was she legally competent to make the decisions she has made?’ is not the same as asking, as Atkins demands, that to respect her autonomy we must ask of the other’s subjective perspective, because competence is set up as an objective, and not a subjective, concept. It is clear from Atkins’s argument, and her example, that it is directed to, and if deployed might have been of much more assistance in, cases such as Bland and Re F, or even in B’s case, if the court had decided that she was incapable of making decisions for herself. Advocating the subjective character of experience in those cases might not only have produced a different result, especially in Re F,⁷¹ but would certainly have given the law a different shape. Butler-Sloss P is, however, adamant that the thesis ought to apply to Ms B’s circumstances: ‘[W]e have to try inadequately to put ourselves into the position of the gravely disabled person and respect the subjective character of experience.’⁷²

If we accept (which at least seems arguable these days) that there are irreducible differences between us that separate us as subjects, then it may be appropriate to ask that we imagine what it would be like for me to be on a ventilator in the same condition, and all other things being equal, as B finds herself. We may wonder, however, what it would be or what it would entail to ask that we imagine — or think — that we are Ms B. Still less is it clear what it adds to what Ms B is saying. In other words, what added value is there to imagining that we are Ms B, rather than just doing as she asks?

Of course, we can understand the good intentions of Butler-Sloss P in adopting this thesis of experiential imagination. Medical staff must try to understand the

---

⁶⁹ Atkins, above n64 at 76.
⁷⁰ Id at 78.
⁷² [2002] 2 All ER 449 at 472. The aspirational character of the thesis is, to some extent, captured in Bob Dylan’s ‘Positively 4th Street’: ‘I wish that for just one time you could stand inside my shoes and just for that one moment I could be you.’
position of individuals such as Ms B and, by so doing, appreciate the values of the chronically ill. Nonetheless, and in addition to the point that this process seems superfluous in a case such as Ms B’s, there is no necessity that imagination need work in any particular manner. To imagine is itself an evaluative process that does not lose this characteristic by pretending to think we are someone else. It is surely possible to come to either conclusion (acceptance/continuation or refusal/withdrawal of medical treatment) after taking into account a variety of factors that we consider are relevant when we try to imagine that we are the ill person in their circumstances. Thus, leaving aside the fundamental problems that imagination poses, let us say I imagine myself to be individual X who is unconscious on a ventilator. I come to the conclusion that she would wish to continue with ventilation and, therefore, continue to provide ventilation. A fairly unremarkable conclusion, one may think. But is the point not that it says more about the conscious individuals who pretend to be X than it does about X in a position where values mean nothing to him or her? We cannot entirely leave our values and identities behind even when we pretend to be someone else. In other words, it is not irrelevant to acknowledge that we are always trying to understand ourselves even when we are attempting to imagine not being ourselves, but someone else — just like medical staff who, in a similar situation, have to contend with their identities as members of the medical profession and all the values that that involves. The main point is, though, that this is another mechanism by which the assessment of individuals occurs in an attempt to respect their autonomy more realistically. This evaluative technique, therefore, helps us to understand a further element of the legal practices surrounding autonomy in medical law.

Let us summarise a couple of the main points made so far. Medical law’s distinction between the nature of patients’ decisions, and their mental capacity to make those decisions, corresponds to the emphasis it places on the division between autonomy (in the sense of a right to choose) and legal tests to establish mental capacity. We have argued, however, that this division is untenable when one considers the legal practices, and the tests themselves, more carefully. The result is that the law’s own tests for mental capacity work to confound the purpose for which they were originally designed — that is, to ensure neutrality by establishing a technical procedure that seeks to prevent account being taken of the nature of the patient’s decision and the reasons behind it.73

We have also tried to demonstrate that those tests, and the law’s discussion of patients’ rights to choose generally, are intimately bound up with what we would call law’s “threshold” function — that is, its central task of deciding on the borders that exist between life and death. Given the significance of this function, it is

73 A related point has been made by Jonathan Montgomery in his discussion of religious beliefs in healthcare law. See Jonathan Montgomery, ‘Healthcare Law for a Multi-Faith Society’ in John Murphy (ed), Ethnic Minorities, Their Families and the Law (2000) at 161–179. However, Montgomery does not think there is anything wrong with the legal tests for competence per se. Rather, it is the manner in which law applies these to the facts of specific cases that he believes is problematic. We have argued that both aspects contribute to undermining the law’s claim to neutrality.
possible to witness broader assessments of the individual occurring — something that is consistent with the focus on explanation, character and imagination as techniques of assessment. This situation further works to undermine the distinction law attempts to make between autonomy and tests for mental capacity.

While to point out this aspect of law’s inconsistency is important, it does little to further our understanding of the nature of legal regulation in this area, save to say that law’s tests, and the practices associated with them, contradict its stated objectives. In order to expand our understanding of this regulation, it is necessary to ask three questions: First, why is it that, in relation to medical practice, new situations of life and death have been created in the last forty years or so; secondly, what were some of the reasons for the increased focus on autonomy — in the general sense that we have sought to describe it — in medicine during this same period; and, finally, what have those general changes meant for the nature of legal regulation in the medical law sphere? Finally, we offer some preliminary reflections on these questions in English and, we anticipate, Australian law.

6. Consequences for Law’s Identity

A. The impact of Technological Development

Medical law offers a fascinating and important window through which to view some significant metamorphoses in contemporary society. Re B is a further case that illustrates the complex relationships amongst death, life, dying, living, the individual and law. The increasing importance — and shifts in meaning — of those different factors owe much more to movements external to law than to any nostalgic belief in the instrumentalism of the latter. It is not possible to discuss all the reasons for those transformations here. Rather, we confine ourselves to a brief discussion of the impact that a couple of technological developments have had in transforming what were then traditional understandings of life and death.

Peter Singer provides a useful account of two related developments — the redefinition of death in the 1960s and the emergence of the ventilator — that had significant consequences for what he calls our ‘traditional ethic’ — that is, the sanctity of life principle.74 The invention of the mechanical ventilator or respirator in the 1950s was intended to replace the function of the brain stem (which makes breathing and heartbeat possible) and, thereby, maintain the lives of individuals long enough for them to make a full recovery from whatever physical intrusion they were suffering. While it undoubtedly had this effect for some, others continued to live without recovering consciousness. As this process could continue indefinitely, the question arose as to the value of persisting with ventilation in such circumstances.

In the 1960s, questions surrounding organ transplantation, especially that of the heart, became pertinent to the discussion regarding artificial ventilation. In order for the heart to be transplanted successfully, it needed to be removed as quickly as possible after death. The many irreversibly unconscious individuals receiving

ventilation in hospitals then came to be seen as a potential source of life-saving organs for other individuals awaiting transplant surgery. The problem was that the removal of an unconscious individual’s heart would amount to murdering him or her. It was this obstacle, together with the futility of providing ventilation to such individuals, that led, in 1968, to the recommendation by the Harvard Brain Death Committee that the definition of death be altered from one of the cessation of breathing and circulation to one of irreversible cessation of all brain function.75 Those falling within the latter category — the functioning of whose organs could be maintained artificially by means of ventilation — could therefore be declared dead and provide a source of organs for transplantation to those who needed them.76

Those early technological developments obviously altered not only the scientific definition of death but produced new types of death and different categories of patient — the individual who will recover sufficiently to be removed from the ventilator; the patient who will remain unconscious, even with the assistance of the machine; and those, like Ms B, who, while conscious, will be unlikely to survive without the assistance of artificial ventilation.77 The latter two indicate the effects that developments in medical technology have had on traditional understandings of death and individuals’, and their families’, experiential relationships with it. There is, therefore, a real sense in which advances in modern science have delivered Ms B — like Tony Bland and Stephen Blood before her — not only to the ward of the hospital, but also to the precincts of the court.

B. The Rise of the Patient and His or Her Autonomy

If technological change produced, and continues to produce, novel perceptions of life and death, it also contributed to the emergence of patient autonomy as a central idea during roughly the same period. However, it took the birth of bioethics as an academic discipline in the late 1960s and early 1970s to highlight the connection between the two.

In his work on the history of medical ethics, Albert Jonsen refers to Joseph Fletcher, who, in 1949, ‘affirmed that in all these problems [of medical ethics], it was the patient who had the right to decide what should be done … The bioethics that would emerge some 20 years [later] was built around the concept of the autonomy of the patient’.78 The emergence of bioethics, however, did not occur simply as a result of the call for an increased focus on the views of patients. Jonsen

---

75 For a simple explanation and review, see Christopher Pallis, ABC of Brainstem Death (1983) and ‘Death’ The New Encyclopaedia Britannica (15th ed, 1986).
76 For a trenchant criticism of this development and a survey of equally forceful objections, see Martyn Evans, ‘Against the definition of brain stem death’ in Morgan & Lee (eds), Death Rites, above n22 at 1.
77 There was a 5 per cent chance that she would survive without ventilation if it were to be withdrawn. She would then become a legal cousin to Diane Pretty, whose almost contemporaneous struggle to achieve a subjectively defined ‘good death’ with dignity failed to gain legal sanction, but for different reasons, in the English courts and the European Court of Human Rights.
78 Jonsen, above n41 at 94.
also traces the birth of this new discipline to the intense research agenda being conducted in the United States between 1945 and 1965. As we have indicated, during this period organ transplantation began and the ventilator was invented. The emergence and success of such new operations and technology brought with them, among other things, a pressing set of ethical dilemmas, the most significant of which concerned the nature and meaning of life and death. As discussions relating to such issues proliferated, Jonsen says that ‘[b]ioethics inherited the task of delineating the principles and values that would shape those ethical dimensions’.

That those principles would include patient autonomy seemed not to be in doubt. Given the often brutal experiments conducted by Nazi physicians on non-consenting prisoners in concentration camps during World War II, purportedly in the name of science, it would have been inconceivable if patient autonomy had not formed a crucial part of the emerging discipline of bioethics. This discipline, and the integral place of autonomy within it, was thus born of two very different types of research — one whose intentions were altruistic; the other whose practices and consequences were horrific.

The need for patient autonomy to form the bedrock of the new discipline also reflected deeper anxieties which, while clearly including the new technological developments in scientific medicine, were not confined to these:

The new concerns about medical technology, combined with a growing skepticism in all areas of life about the impartiality of professional experts, summoned a principle unfamiliar to traditional medical ethics but familiar to philosophers: the freedom of persons to judge what is in their benefit without interference from others.

Thus, along with developments in medical technology, it was scepticism of professionals that contributed to the emergence of patient autonomy as a central feature of the new bioethics. Indeed, to a large extent, those very developments gave rise to bioethics itself, and a *symbiotic* relationship seemed to arise between patient autonomy and bioethics — patients needed a powerful voice for their views to be heard, and bioethics required a core principle upon which to found a credible discipline. Thus, as well as progress in medical technology and the existence of a sceptical public, it would seem that the institutional needs of a new discipline — bioethics — were central in placing the patient not only at the heart of medical practice, but at the forefront of debate concerning future developments in scientific medicine. And, as the previous quotation indicates, the emphasis on freedom from interference helped shift the status of patients from those who would readily submit to the wonders of medicine to those of *individuals* who would now confront medicine, and the circumstances it was capable of producing, with their own estimations of what was acceptable.

79 Id at 111.


81 Jonsen, above n41 at 116–117.
C. The Nature of Legal Regulation

It is our argument that points A. and B. of this section are both partially responsible for the approach of the law we see in *Re B*. First, the situation in which Ms B found herself would not have arisen without the ventilator. Its workings produced circumstances in which life and death retained no clear distinction for her. Speaking about her life attached to the ventilator was, as we pointed out earlier, also to speak about her death. Identifying the threshold between the two seemed no easy task for her. Similarly, this question of the threshold between life and death becomes the principal axis around which law must operate. And, while Butler-Sloss P was unwilling to view the court’s function in such terms, we have already seen how some senior members of the judiciary are not only acknowledging this role, but claiming law’s pre-eminence where disputes regarding this threshold emerge. Deciding what the boundaries of life as such are is becoming one of the core functions of medical law.

Secondly, there is no doubt that autonomy is a central idea in medical law. But we would argue that, despite the judiciary’s protestations to the contrary, it should not be thought of today primarily as the freedom from external interference that Jonsen described its meaning as having in the early days of bioethics. Rather, the manner in which we have described the operation of autonomy in *Re B*, together with the various practices that have emerged around it, suggests that autonomy is now more widely connected to an obsession with the individual generally. It is no longer simply about patients’ rights to choose; it is a notion that, in many ways, describes a whole subject — medical law — and how that subject ought to be studied and practised. Moreover, there is a real sense in which the individual situations of those like Ms B — we are thinking of the cases of *Bland, Blood, Pretty* and (we might say) so on — are very public affairs. As those situations pass from the hospital to the court, they never escape institutional control and management. The critical point, though, is that this control occurs through a focus on the individual and, in particular, through the filter of the legal emphasis on autonomy.

This combination of life, death, and autonomy has implications for the nature of legal regulation here. In his discussion of bio-power, Michel Foucault says of law:

> I do not mean to say that the law fades into the background or that the institutions of justice tend to disappear, but rather that the law operates more and more as a norm, and that the judicial institution is increasingly incorporated into a continuum of apparatuses (medical, administrative, and so on) whose functions are for the most part regulatory. *A normalizing society is the historical outcome of a technology of power centered on life.* [Emphasis added.]

Foucault’s discussion of bio-power centred, of course, on the methods used to ‘correct’ lives in certain institutions (disciplinary mechanisms) and, later, the

---

82 Michel Foucault, *The Will to Knowledge: The History of Sexuality, Volume 1* (Robert Hurley, trans, 1978) at 144.
conduct of lives in the population generally (bio-politics). The proliferation of medical, and now human genetic, technologies means that the biological aspects of life are becoming more important as bases from which evaluations about people and the ways in which they live can be conducted. Thus, as Paul Rabinow has recently pointed out, Foucault’s bio-power needs contemporary elaboration, to encompass the very building blocks of life itself — DNA. So, even if Foucault’s notion of life requires expansion, his reference to law as a norm appears now particularly useful as a way of understanding how medical law operates in cases involving decisions about life and death.

Re B illustrates that the combination of life, death, and autonomy does not result in a polarised battle between the sanctity of life principle and the freedom to live one’s own life as one sees fit. Decisions are not made by simply choosing one as opposed to the other. Rather, and as the significance of character, explanation, imagination, and tests for mental capacity indicates, the law operates in a much more subtle manner, regulating what it considers to be the threshold between life and death by constructing norms through patients. Through the power to maintain life technologically, and with the increased focus on the individual, those are the techniques that have arisen in medical law to allow it to be described as a normalising practice. If life is to be allowed to end (‘this grave decision’, as Butler-Sloss P refers to it in Re B) and the individual has priority, the search for norms must occur through that individual. The question that Foucault used to describe the essence of the modern criminal tribunal — ‘Who are you?’ — is becoming increasingly useful as a means of capturing a fundamental aspect of the medical law arena.

Case Commentary

A Prelude to the Demise of Teoh: The High Court Decision in Re Minister for Immigration and Multicultural Affairs; Ex parte Lam

WENDY LACEY*

Abstract

The authority of the High Court’s 1995 decision in Teoh must now be examined in light of the High Court’s 2003 decision in Lam. In the latter case, the minority view of McHugh J expressed in Teoh appears to have gained ascendancy among a majority of the present High Court judges. Though the critique on Teoh offered in Lam is contained in obiter comments only, the principle that ratification of a treaty may give rise to a legitimate expectation that administrative decision-makers will act in accordance with the terms of the treaty, has been seriously called into question.

In this paper, each of the separate judgments provided in Lam are examined in detail. The argument is presented that the reasoning employed by the Court is open to criticism on several grounds. These include the introduction of private law concepts into the public law sphere, and the failure to distinguish between questions going to the existence of a breach of procedural fairness, and questions that pertain to the discretionary power to grant or deny a remedy. Though an investigation of the continuing role that legitimate expectations may play in administrative law should be welcomed, it is argued that the reasoning adopted in Lam is flawed, and is likely to create only further confusion surrounding legitimate expectations generally.

* Lecturer in Law, University of Adelaide. This paper was originally presented as a seminar paper to the Comparative Administrative Law Class, at the Faculty of Law, University of Tasmania, 17 April 2003. The author would like to acknowledge the valuable comments made on that paper by Sir Anthony Mason, Dr Steven Churches, Mr Chris Finn, Mr Rick Snell, and Professor Mark Aronson. However, the views expressed herein solely reflect those of the author.
1. **Introduction**

The latest High Court decision to address the subject of legitimate expectation, including the court’s previous discussion in *Teoh*,¹ is the case of *Re Minister for Immigration and Multicultural Affairs; Ex parte Lam*.² The applicant, Mr Lam, had had his Transitional (Permanent) Visa cancelled in early 2001 and his deportation ordered, following a conviction for serious offences. Like many in this situation, Mr Lam had established a family during his time in Australia, and his deportation would have either entailed relocating his Australian-born children to another country, or Mr Lam being separated from them. While these facts raise important issues regarding the government’s policy to deport in such instances,³ the case itself raised no significant or novel legal question. Neither *Teoh*, nor the more recent and controversial English case, *R v North and East Devon Health Authority; Ex parte Coughlan*,⁴ were relied upon in argument by the applicant. The latter case, which identified circumstances where a decision-maker would be required to offer substantive as opposed to merely procedural fairness, has provoked strong criticism in Australia. Due to Australia’s constitutional framework, however, the case is unlikely to gain a foothold in Australian administrative law. For this reason, and in order to provide a specific framework for analysis in the present context, the *Coughlan* decision is only considered in the extent to which it was referred to in the separate judgments in *Lam*. The focus of this paper is therefore directed specifically at the court’s treatment of *Teoh* and legitimate expectation.

Notwithstanding that neither *Teoh* nor *Coughlan* were relied upon in argument, the decision in *Lam* is extraordinary for a number of reasons. It foretells of significant changes, or more precisely limitations, to the role of legitimate expectations in Australian administrative law, as well as to the doctrine in *Teoh*. The case is also noteworthy for the fact that nearly two thirds of its length is devoted to issues not directly raised by the facts of the case and which consequently must be taken to be obiter statements. Yet, within those obiter statements can be discerned an intent to fundamentally disturb the role of legitimate expectations, as well as its application as articulated in the *Teoh* decision.

While the sheer length of obiter comments is concerning in itself, this prelude to the demise of *Teoh* and the confinement of legitimate expectations generally can be shown to rest on the application of flawed reasoning ill-suited to the public law context. In particular, private law concepts such as estoppel have been employed by the court in a manner that lends itself to an assessment of substantive (as opposed to procedural) unfairness. For reasons concerning both the legitimacy of the court’s role in conducting judicial review, as well as consistency in administrative decision-making generally, it is argued that this trend should be

---

¹ (1995) 183 CLR 273 (hereinafter *Teoh*).
² (2003) 195 ALR 502 (hereinafter *Lam*).
³ These issues have arisen in a number of recent cases involving similar factual scenarios: see, for example, *Ayan v Minister for Immigration and Multicultural and Indigenous Affairs* (2003) 196 ALR 332.
⁴ [2001] QB 213; [2000] 3 All ER 850 (hereinafter *Coughlan*).
resisted. In addition, it is submitted that the tendency of the court in Lam to conflate the question of whether there has occurred a breach of procedural fairness with the question of whether a discretionary remedy should be granted has only added to the confusion surrounding Lam.

2. **Historical Background: The Decision in Teoh**

Like Mr Lam, the applicant in Teoh had been subject to a deportation order following a criminal conviction for importing heroin into Australia. The facts and proceedings of Teoh are outlined extensively by commentators elsewhere, and need not be reproduced here. It may be stated, however, that Mr Teoh was essentially the responsible parent for seven Australian-born children (including Mr Teoh’s biological and step-children). What is important in the present context is to revisit the legal principles established in Teoh’s case, and in particular the dissenting view of McHugh J in Teoh. Consequently, the separate judgments of Toohey and Gaudron JJ will not be addressed in the present paper.

At one level, the majority of the High Court simply rearticulated the existing principles regarding the legitimate use that may be made of unincorporated international treaties in Australian law. Thus, as Mason CJ and Deane J stated:

> [T]he fact that the Convention has not been incorporated into Australian law does not mean that its ratification holds no significance for Australian law. Where a statute or subordinate legislation is ambiguous, the courts should favour that construction which accords with Australia’s obligations under a treaty or international convention to which Australia is a party, at least in those cases in which the legislation is enacted after, or in contemplation of, entry into, or ratification of, the relevant international instrument. That is because Parliament, prima facie, intends to give effect to Australia’s obligations under international law.

It is accepted that a statute is to be interpreted and applied, as far as its language permits, so that it is in conformity and not in conflict with the established rules of international law.

---


6 Toohey J was largely in agreement with Mason CJ and Deane J. Gaudron J took a different perspective, based on the fact that a legitimate expectation was not even needed. For Gaudron J, the common law contained a requirement that in all matters affecting children their welfare would be a primary consideration. For a discussion of these judgments see Allars, above n5.

7 Teoh, above n1 at 287.

8 Chu Kheng Lim v Minister for Immigration (1992) 176 CLR 1 at 38.

9 Polites v The Commonwealth (1945) 70 CLR 60 at 68–69, 77, 80–81.
Their Honours also referred to the accepted use that may be made of international instruments in the development of the common law.10

The provisions of an international convention to which Australia is a party, especially one which declares universal fundamental rights, may be used by the courts as a legitimate guide in developing the common law.11 But the courts should act in this fashion with due circumspection when the Parliament itself has not seen fit to incorporate the provisions of a convention into our domestic law. Judicial development of the common law must not be seen as a backdoor means of importing an unincorporated convention into Australian law.

In these respects, the majority judgment was uncontroversial. The principal legal controversy concerned the question of whether ‘Australia’s ratification of the Convention [on the Rights of the Child] … [could] give rise to a legitimate expectation that the decision-maker will exercise that discretion in conformity with the terms of the Convention.’12 The fact that the majority of the court, including Mason CJ and Deane J (together with Toohey J who delivered a separate judgment), answered this question in the affirmative, was what created the controversy surrounding the decision.13 In addressing the question posed, the joint judgment offered the following reasoning:

[R]atification by Australia of an international convention is not to be dismissed as a merely platitudinous or ineffectual act,14 particularly when the instrument evidences internationally accepted standards to be applied by courts and administrative authorities in dealing with basic human rights affecting the family and children. Rather, ratification of a convention is a positive statement by the executive government of this country to the world and to the Australian people that the executive government and its agencies will act in accordance with the convention. That positive statement is an adequate foundation for a legitimate expectation, absent statutory or executive indications to the contrary, that administrative decision-makers will act in conformity with the Convention15 and treat the best interests of the children as “a primary consideration”. [Emphasis added.]

In relation to the expectation identified in Teoh, the fact that an individual may not have been aware of the Convention’s existence would not preclude the expectation

10 Teoh, above n1 at 288.
11 Mabo v Queensland [No 2] (1992) 175 CLR 1 at 42 (Brennan J, with whom Mason CJ and McHugh J agreed); Dietrich v The Queen (1992) 177 CLR 292 at 321 (Brennan J); at 360 (Toohey J); Jago v District Court (NSW) (1988) 12 NSWLR 558 at 569 (Kirby P); Derbyshire County Council v Times Newspapers Ltd [1992] QB 770.
12 Teoh, above n1 at 288.
13 For an overview of the political and legal response to the court’s decision, including two executive statements, three federal bills, one state Act, and various cases at the Federal Court level, see Wendy Lacey, ‘In the Wake of Teoh: Finding an Appropriate Government Response’ (2001) 29 Fed LR 219–240.
14 See Minister for Foreign Affairs and Trade v Magno (1992) 37 FCR 298 at 343; Tavita v Minister of Immigration [1994] 2 NZLR 257 at 266.
from arising. For Mason CJ and Deane J, it was enough that the expectation was ‘reasonable in the sense that there are adequate materials to support it’. The effect of the legitimate expectation was not to create a binding rule of law, whereby a decision-maker was compelled to act in conformity with the Convention. Such a result would have involved the incorporation of the Convention ‘by the backdoor’ and would have linked the effect of the legitimate expectation with substantive outcome, rather than with the procedures to be adopted.

This distinction is critical in administrative law, and stems from the separation of powers in Australia. Judicial review of administrative action by the courts is limited to consideration of the legalities of decision-making, rather than the actual merits of the case — which is an administrative function to be performed by the executive. Indeed, it is by virtue of the existence of a constitutional separation of powers in Australia that the decision in Coughlan stands at odds with Australia’s constitutional framework, and is unlikely to be followed here.

The actual effect of the legitimate expectation identified in Teoh was not of this class of decision, however. It did not involve a development of the law in conflict with the constitutional separation of powers. The effect of the legitimate expectation in Teoh’s case was limited to a procedural guarantee, which was itself premised upon the right of the decision-maker to act inconsistently with a treaty obligation. Where a decision-maker proposed to decide a matter inconsistent with a legitimate expectation, procedural fairness would require that any persons affected be given notice and an adequate opportunity of presenting a case against such a course of action.

However, in the High Court’s decision in Teoh, Justice McHugh delivered a strong dissenting judgment, that treated the legitimate expectation identified by the majority as one involving the substantive protection of the treaty, rather than being concerned with procedural requirements. He began his discussion of the issue with the following statements:

In my opinion, no legitimate expectation arose in this case because: (1) the doctrine of legitimate expectations is concerned with procedural fairness and imposes no obligation on a decision-maker to give substantive protection to any right, benefit, privilege or matter that is the subject of a legitimate expectation; (2) the doctrine of legitimate expectations does not require a decision-maker to inform a person affected by a decision that he or she will not apply a rule when the decision-maker is not bound and has given no undertaking to apply that rule; (3) the ratification of the Convention did not give rise to any legitimate expectation that an application for resident status would be decided in accordance with Art 3.

16 Teoh, above n1 at 291.
17 Ibid.
18 As Brennan J stated in Attorney-General (NSW) v Quin (1990) 170 CLR 1 at 36: ‘The merits of administrative action, to the extent that they can be distinguished from legality, are for the repository of the relevant power and, subject to political control, for the repository alone.’ His Honour continued at 38: ‘If judicial review were to trespass on the merits of the exercise of administrative power, it would put its own legitimacy at risk.’
19 Teoh, above n1 at 291–292.
20 Id at 305–306.
Obviously the point from which McHugh J departed was, therefore, entirely different from that of the majority. From the very outset then, they were approaching the question from different perspectives. In his judgment, McHugh J also questioned several aspects of legitimate expectations. First, his Honour questioned the practical role for the doctrine following the decision in *Kioa v West*, where the High Court had adopted a broad approach to the types of interests sufficient to enliven procedural fairness requirements. While this is an accurate analysis of the position since *Kioa* regarding what is referred to as the ‘threshold test’, it is submitted that the doctrine of legitimate expectations must still play a role in determining the content of procedural fairness requirements in any given case.

The other criticism contained in McHugh J’s judgment, concerned the objective nature of the expectation in *Teoh*. For McHugh J, the subjective state of mind of an individual affected by a decision should be treated as relevant to whether the expectation is ‘objectively reasonable’, as the following statement indicates:

> It must be an expectation that is *objectively reasonable* for a person in the position of the claimant. But that does not mean that the state of mind of the person concerned is irrelevant. … If a person does not have an expectation that he or she will enjoy a benefit or privilege or that a particular state of affairs will continue, no disappointment or injustice is suffered by that person if that benefit or privilege is discontinued. A person cannot lose an expectation that he or she does not hold. [Original emphasis.]

Clearly, the reasoning in McHugh J’s dissenting judgment stands directly at odds with that of the majority. However, following changes in the membership of the High Court since 1995, and with McHugh J being the only judge remaining from the *Teoh* decision, *Lam* provided an opportunity to again test the reasoning of the majority in that case. Though, as has been pointed out, *Teoh* was not relied upon in argument (as no legitimate expectation was premised upon the ratification of a treaty), in 2003 McHugh J found himself on a bench and among judges sympathetic to the views he had earlier expressed in dissent. It is against this background, that the decision in *Lam* must be considered.

### 3. The Factual Background in Lam

Mr Lam, the applicant, was born in Vietnam but had arrived in Australia as a 13 year old refugee in 1983. On arrival he was granted a Transitional (Permanent) Visa, which was cancelled on 23 January 2001 pursuant to a decision of the Minister under s501(2) of the *Migration Act* 1958 (Cth). The applicant had been convicted of several criminal offences, the most serious of which had been trafficking in heroin for which Mr Lam had been sentenced to eight years imprisonment. As a consequence, the applicant had failed to meet the character test contained in s501(6) of the Act, and became liable to deportation.

---

21 *(1985) 159 CLR 550* (hereinafter *Kioa*).
22 *Teoh*, above n1 at 314.
The proceedings before the High Court involved an application for orders of certiorari and prohibition under s75(v) of the Constitution, to quash the Minister’s decision and to prevent the Minister from taking steps to deport Mr Lam. The major argument presented by counsel for the applicant involved an allegation of want of procedural fairness, based on the failure of a departmental officer to follow an announced procedure. The relevant facts were as follows.

Mr Lam, although unmarried, was the father of two girls, both of whom had been born in Australia in 1989 and 1993 respectively, and who were Australian citizens. Mr Lam had been estranged from the girls’ mother for some time, and the girls had been living with relatives for much of Mr Lam’s period of imprisonment. Mr Lam had entered into another relationship with an Australian woman and was engaged to be married.

In September 2000 an officer of the Department of Immigration and Multicultural Affairs had written to Mr Lam indicating that his visa may be liable to cancellation under s501 of the *Migration Act* 1958 (Cth). In the letter, details of the legislation had been set out, an opportunity for Mr Lam to comment prior to any decision of the Minister was given, and a list of relevant matters to be taken into account in making a decision under s501 was set out. Those matters included ‘the best interests of any children with whom you have an involvement’.

Mr Lam exercised his right to comment on the matters relevant to the decision to be made, in what Gleeson CJ referred to as a lengthy submission ‘obviously prepared with skilled assistance’.23 The applicant’s letter was dated 30 October 2000. On the matter relating to his children, the applicant provided information regarding his daughters and their current circumstances, and advanced arguments as to why their interests required that he not be deported. These arguments included the fact that the girls had no contact with people from Vietnam, that they were happily settled in Australia, that he planned to marry upon his release from prison, and that the children would have to be cared for by the state were he to be deported.24 Annexed to Mr Lam’s letter, were letters from both his fiancée and from the children’s carers, represented in this case by a letter from Ms Tran. The letter from the carer supported the facts raised in Mr Lam’s letter regarding the children’s circumstances, and his arguments related to his deportation. In particular, the letter indicated that the welfare of the children would be best served in the long term if the children were cared for by Mr Lam and his fiancée.

On 7 November 2000, an officer of the ‘Character Assessment Unit’ within the Department wrote to Mr Lam, stating the following:25

The United Nations Convention on the Rights of the Child provides that in all actions concerning children, the best interests of the children shall be a primary consideration.

---

23 Lam, above n2 at [7] (Gleeson CJ).
24 Ibid.
25 Id at [9].
Would you therefore kindly provide the full name, address and telephone number of the children’s carers. The Department wishes to contact them in order to assess your relationship with the children, and the possible effects on them of a decision to cancel your visa.

Would you please provide the full contact details of the mother of the children as well.

Though Mr Lam replied with the appropriate information, the Department chose not to take any steps to contact the children’s carers, opting to prepare a document for the Minister on the information already in their possession. No problems arose in relation to the contents of this document, only in relation to the Department’s failure to contact the carer of the children. Because of this it was argued that procedural fairness was not provided, as Mr Lam had not been informed of the decision not to contact the carer. This was argued despite the fact that no evidence was directly led indicating that Mr Lam ‘was misled into taking or failing to take some step, or deprived of an opportunity to advance his case in some way’.26

The transcripts of the hearing before the High Court indicate that evidence concerning the language background of the children and of their capacity to integrate into Vietnamese society could have been introduced.27 This evidence was not directly used to argue a denial of procedural fairness. Rather, it was left open for the court to find that the absence of an opportunity to present further evidence was insufficient. The basic premise of the applicant’s case was that Mr Lam was denied procedural fairness, as he had a legitimate expectation created by the letter of 7 November that was not fulfilled. While the decision-maker was not bound to comply in substance with that expectation, the argument presented was that Mr Lam should have been notified of the decision not to follow through with the representation made to Mr Lam, affording him an opportunity to respond.28 The argument was rejected by the High Court, and the reasoning in each of the separate judgments is outlined below.

4. **Chief Justice Gleeson**

Chief Justice Gleeson accepted that in some circumstances procedural unfairness will result where an administrative decision-maker states to a person affected an intention to take a procedural step and fails to do so without warning the person affected of the change in intention. His Honour rejected, however, that procedural unfairness would result in all cases where such a change was made from the stated intention. His reasoning was largely premised on the discretionary nature of the remedies sought, including certiorari and prohibition, as well as the actual requirement of fairness. According to Gleeson CJ:29

> There are undoubtedly circumstances in which the failure of an administrative decision-maker to adhere to a statement of intention as to the procedure to be

---

26 Id at [20].
27 See *Re MIMA; Ex parte Lam* B33/2001 (24 June 2002) High Court of Australia Transcripts.
28 *Lam*, above n2 at [23].
29 Id at [25].
followed will result in unfairness and will justify judicial intervention to quash the
decision; but for the present applicant to succeed it would be necessary to
conclude that such a result will follow in all circumstances. That cannot be
correct. To begin with, it overlooks the discretionary nature of the remedies of
certiorari and prohibition.30 And, in any event, it requires the concept of
legitimate expectation to carry more weight than it will bear. If such a proposition
were accepted, it would elevate judicial review of administrative action to a level
of high and arid technicality.

In approaching the present case, Gleeson CJ referred to the Privy Council’s
decision in Attorney-General (Hong Kong) v Ng Yuen Shiu,31 and the decision of
Dawson J in Attorney-General (NSW) v Quin.32 For Gleeson CJ, those two
decisions stood for the proposition that ‘when a public authority promises that a
particular procedure will be followed in making a decision, fairness may
require that the public authority be held to its promise’.33 What needed to be shown before
the remedies of certiorari and prohibition would be granted was, for Gleeson CJ,
actual unfairness rather than ‘[a mere] departure from a representation’.34 As his
Honour emphatically stated, ‘[n]ot every departure from a stated intention
necessarily involves unfairness, even if it defeats an expectation’.35 Gleeson CJ’s
concept of fairness directly correlated with what the applicant was entitled to
expect on the part of the decision-maker. He considered that fairness did not
require departmental officers to contact Mr Lam prior to changing their minds
about contacting the girls’ carer.36 This conclusion was partially based on
supposition from the facts, as the following extract shows:37

[T]here could have been a number of reasons why they might change their plans,
without necessarily having to inform the applicant. Let it be supposed, as may
well be the case, that they changed their minds because they realised that they had
already heard from Ms Tran, they did not doubt what she had to say, and it was
unlikely that there was anything she could usefully add to what had already been
said. … The applicant does not seek to show that such a view was not reasonably
open. I do not accept that it would have been reasonable to expect the department
to write to the applicant if for any reason there was a change of plan about
contacting Ms Tran.

In his judgment Gleeson CJ clearly articulates the concept of ‘detriment’ upon
which his requirement of ‘unfairness’ rests.38 In addition, the existence of a
subjective expectation was employed to indicate the specific representation upon

30 Re Refugee Review Tribunal; Ex parte Aala (2000) 204 CLR 82; 176 ALR 219.
32 Above n18 at 56–57.
33 Lam, above n2 at [33] (Gleeson CJ).
34 Id at [34].
35 Ibid.
36 Id at [35].
37 Ibid.
38 Id at [36]–[37].
which a person relies to their detriment, and which consequently results in unfairness. According to Gleeson CJ:

[(I]t is the existence of a subjective expectation, and reliance, that results in unfairness. Fairness is not an abstract concept. It is essentially practical. Whether one talks in terms of procedural fairness or natural justice, the concept of the law is to avoid practical injustice.

No practical injustice has been shown. The applicant lost no opportunity to advance his case. He did not rely to his disadvantage on the statement of intention. It has not been shown that there was procedural unfairness.

The notions of reliance and detriment to which his Honour refers are taken from the doctrine, more appropriately applied in private law, of estoppel. In addition, the reference to ‘subjective expectation’ highlights the individual applicant’s actual expectation over any objective or constructive expectation. This emphasis narrows the consequences of representations or undertakings by administrative decision-makers to those that directly and personally relate to the applicant’s case. It ‘privatises’ the notion of expectation in administrative law.

The problem with employing private law concepts within a public law setting rests directly with the limitations of judicial review. The fact that consideration by a court into the actual detriment suffered by an individual tends towards a consideration of the actual merits of a case, offers a significant caution against its use in cases involving legitimate expectation. Indeed, as Aronson and Dyer have observed, there has been a distinct avoidance of such inquiries in cases involving legitimate expectations. While, like Gleeson CJ, they treat the primary consideration as being what fairness requires in the circumstances, rather than actual compliance with an undertaking or representation, for Aronson and Dyer, the presence of reliance and detriment may affect the content of procedural fairness, but it is not critical to whether procedural fairness was denied. This issue was aptly raised in the hearing before the High Court in a passage between Mr Walker (Counsel for the Applicant) and Gleeson CJ:

GLEESON CJ: Well, a possible point of view — to paraphrase a proposition from the law of contract — is that an unrelied on representation is a thing writ in water.

MR WALKER: Yes, that is a possible point of view, your Honour. This case certainly raises for consideration whether or not there are sufficient analogies with representational causes of action in, most species of estoppel, so as to give rise to a requirement of what I will call actual psychological belief, on the basis

39 Id at [37].
40 Id at [37]–[88].
41 Though, detriment in relation to the ‘procedures’ involved, rather than the actual substantive outcome, would not necessarily tend towards such ‘merits’ considerations.
43 Ibid.
of which specified acts or omissions took place. In our submission, it would be a
misstep and error for this Court to confine administrative law, in relation to
procedural fairness aspects of natural justice, in that fashion.44

The role of estoppel in the context of administrative law raises particular problems. As Gummow J stated in *Minister for Immigration, Local Government and Ethnic Affairs v Kurtovic*, 'there are considerable difficulties in the way of propounding an estoppel against the exercise of an administrative discretion'.45 As a doctrine that has evolved in private law, estoppel has a limited application in administrative law given that it cannot operate to prevent or limit the exercise of a public duty or discretion granted under statute.46 Mason CJ extended this principle to all public powers whether they are sourced in statute, prerogative or the common law in *Attorney-General (NSW) v Quin*.47 Gummow J, however, raised the possibility of a distinction between powers or discretions which fall within the object of the statute (the ‘policy’ basis of the power or discretion), and operational decisions which relate to the implementation of the core statutory powers and discretions — the latter perhaps attracting the application of estoppel.48 His Honour made the following qualified statement on this point:

Where the public authority makes representations in the course of implementation of a decision arrived at by the exercise of its discretion, then usually there will not be an objection to the application of a private law doctrine of promissory estoppel. It must, however, be recognised that it may be difficult, in a given case, to draw a line between that which involves discretion and that which is merely “operational”.49

The analysis of Gummow J in *Kurtovic* of the role of estoppel in administrative law highlights, therefore, the potential for certain ‘operational’ public powers to attract the private law doctrine of estoppel. While this of itself may be problematic, given the difficulty in identifying ‘operational’ powers, the need to establish detrimental reliance raises even more problems. Gummow J was of the view that detrimental reliance would be just as relevant in public law as in private law. In doing so, he rejected the competing view that ‘the citizen is entitled to expect public authorities and those they employ to deal with the public will keep their word and act within their authority, so that detrimental reliance is not necessary’.50 However, as the analysis below outlines, private law notions such as detrimental reliance raise considerable problems when applied in an administrative law context.

44 *Re MIMA* Transcripts, above n27.
45 (1990) 92 ALR 93 at 108 (Gummow J).
46 Ibid at 109.
47 Above, n18 at 18.
48 (1990) 92 ALR 93 at 116-117.
49 Ibid at 116.
What Gleeson CJ does in the Lam decision is utilise private law concepts in establishing a threshold for breaches of procedural fairness where a decision-maker makes a representation regarding procedure. In doing so, he introduces a concept of substantive ‘unfairness’ into the question of whether a breach has occurred, rather than placing it within the question of whether the discretionary remedy should, or should not, be granted. In taking this perspective in Gleeson CJ’s judgment, one is not asserting that a decision-maker is bound to act consistently with the representation made to an individual regarding procedure. It is simply to accept that such a representation will affect the content of procedural fairness, and what the hearing rule will entail in the circumstances. Taking Gleeson CJ’s perspective, the representation will only affect the content of procedural fairness to the extent that a breach occurs, and a breach will only ever occur where the individual to whom the representation was made has relied upon the representation to his or her detriment. In Lam’s case, no detriment had occurred despite the fact that Mr Lam had been denied an opportunity to reconsider his position and whether further evidence should have been submitted on his behalf. As already mentioned, further evidence was available to be tendered. However, it appears that Gleeson CJ was convinced that any further evidence would not have advanced Mr Lam’s case.

There are several issues of concern in relation to Gleeson CJ’s approach. First, the court is drawn into considerations of reliance and detriment, which have the potential to draw the court into issues touching upon the merits. Secondly, the court is forced to consider, in relation to whether a breach has occurred, the issue of ‘unfairness’ in actual outcome, which is more appropriately taken as a matter hinging upon the discretion to grant or deny a remedy. Thirdly, by setting such a high threshold for breaches of procedural fairness where a decision-maker makes a representation to an individual regarding procedure, Gleeson CJ gives priority to administrative efficiency over good administrative practices. In cases such as Lam, what matters is that an individual be informed of any departure from a stated representation as to procedure in order to prevent any reliance upon it in the first place. The decision-maker need not comply with the representation, but should inform the individual that a change in stated procedure is to occur, which would give the individual an opportunity to reconsider their case and to take further action in support of their case if desired. Being denied this opportunity is what constitutes the breach of procedural fairness (rather than departure from the stated intention), whereas detrimental reliance will simply exacerbate any breach.

While Gleeson CJ’s judgment may be criticised on this basis, he refrained from questioning the High Court’s previous decision in Teoh. He also accepted the fact that counsel had not ‘explicitly’ argued for any substantive protection arising from legitimate expectations, and consequently, limited his comments on the decision of the English Court of Appeal in R v North and East Devon Health Authority; Ex parte Coughlan. Gleeson CJ did, however, raise the idea that ‘substantive

51 See note 28 and accompanying text.

52 Coughlan, above n4.
expectations’, in the sense that a decision-maker is bound to act in a manner consistent with a legitimate expectation, could not be as easily found in Australia as in the United Kingdom. In considering that the applicant in Lam had invoked the jurisdiction under s75(v) of the Constitution in seeking judicial review, his Honour made the following comments:

That jurisdiction exists to ensure that exercises of power by officers of the Commonwealth conform to law. The scope of the concept of abuse of power, insofar as it may embrace substantive unfairness of the kind considered in Coughlan, and its relation to s 75(v) of the Constitution, was not the subject of argument, and does not arise for decision. It is a subject that may involve large questions as to the relations between the executive and judicial branches of government.53

While Gleeson CJ felt that the applicant’s argument came close to approaching ‘an attempt to convert a procedural expectation into something substantive’, 54 his Honour accepted that it did not go so far as to attract criticism on that basis. The applicant’s argument had been qualified by the concession that a decision not to follow the stated procedure would be acceptable provided that the applicant was notified of that change. Consequently, Gleeson CJ held that it did not incorrectly treat the legitimate expectation as requiring a decision-maker to act in a particular way — an approach that was ‘tantamount to treating it as a rule of law’. 55 Such an approach had been explicitly rejected by Mason CJ and Deane J in Teoh.56 In respect of Mr Lam’s application, Gleeson CJ dismissed the case with costs.

5. Justices McHugh and Gummow

The tenor of the joint judgment delivered by McHugh and Gummow JJ is indicated very early in the decision when their Honours state that the notion of legitimate expectations, while serving the purpose of standing rules well, ‘remains of limited utility elsewhere’. 57 However, the observation that the argument of the applicant ‘invites attention to the doctrine of “legitimate expectation” gives little indication of the full extent of their Honour’s critique of the concept and its application in Teoh that follows. To find a clear indication to that effect one need only peruse the transcript of the hearing before the High Court. During questioning, counsel for both parties were drawn into considering cases not even relied upon, nor directly relevant to the proceedings. Included were the decisions in both Teoh and Coughlan. While McHugh J’s strong dissent in the Teoh decision already offered a clear indication of his views on the subject, Gummow J’s questions had offered a unique insight into his own thoughts on the subject of legitimate expectation. After drawing Mr Gageler (counsel for the respondent), into a discussion on the

54 Ibid.
55 Ibid.
56 Teoh, above n1 at 291.
57 (2003) 195 ALR 502 at [47].
A. Coughlan and Substantive Protection

Despite the fact that no reliance was placed on the decision in *Coughlan*, and no attempt was made by the applicant to argue for any substantive protection to be granted under the rubric of legitimate expectation, McHugh and Gummow JJ addressed those matters in detail. Their Honours summarised the English developments as follows:

The doctrine of “legitimate expectation” has been developed in England so as to extend to an expectation that the benefit in question will be provided or, if already conferred, will not be withdrawn or that a threatened disadvantage or disability will not be imposed. This gives the doctrine a substantive, as distinct from procedural, operation.

In referring to the decision of Mason CJ in *Attorney-General (NSW) v Quin*, where it was held that natural justice did not entitle an individual to substantive protection in the form of an order requiring a decision-maker to act in a particular way, their Honours stated:

That remains the position in this Court and nothing in this judgment should be taken as encouragement to disturb it by adoption of recent developments in English law with respect to substantive benefits or outcomes.

In considering *Coughlan*, their Honours noted the link between the doctrine of legitimate expectation and substantive benefits, with the notion of unfairness amounting to an “abuse of power”, and questioned its application within an Australian constitutional context:

The notion of “abuse of power” applied in *Coughlan* appears to be concerned with the judicial supervision of administrative decision-making by the applications of certain minimum standards now identified by the English common law. These standards fix upon the quality of the decision-making and thus the merits of the outcome. As was indicated in *Coughlan* itself, this represents an attempted assimilation into the English common law of doctrines derived from European civilian systems.

In distinguishing the English development with the Australian setting, their Honours referred to decisions in both New Zealand and Canada on the matter. In particular, they offered support for the decisions of the Canadian Supreme Court.

---

58 *Re MIMA; Ex parte Lam* B33/2001 (24 June 2002) High Court of Australia Transcripts.
59 *Coughlan*, above n4.
60 (2003) 195 ALR 502 at [66].
61 Above n18 at 22–23.
63 Id at [73].
in Baker v Minister of Citizenship and Immigration64 and Mount Sinai Hospital Center v Quebec (Minister of Health and Social Services).65 Quoting from the latter case, their Honours extracted the following comments from the decision of Binnie J:

> It thus appears that the English doctrine of legitimate expectation has developed into a comprehensive code that embraces the full gamut of administrative relief from procedural fairness at the low end through ‘enhanced’ procedural fairness based on conduct, thence onwards to estoppel (though it is not to be called that) including substantive relief at the high end, ie, the end representing the greatest intrusion by the courts into public administration ….

In ranging over such a vast territory under the banner of ‘fairness’, it is inevitable that sub-classifications must be made to differentiate the situations which warrant highly intrusive relief from those which do not. Many of the English cases on legitimate expectations relied on by the respondents, at the low end, would fit comfortably within our principles of procedural fairness. At the high end they represent a level of judicial intervention in government policy that our courts, to date, have considered inappropriate in the absence of a successful challenge under the Canadian Charter of Rights and Freedoms.66

What is astounding about their Honours’ judgment is the length to which they felt it necessary to address an issue that was not relied upon by the applicant, and indeed, which was explicitly rejected by Counsel in argument. The extent of obiter comments made within their Honours’ judgment was, however, not limited to the developments most noticeably observed in the Coughlan decision. The decision in Teoh did not escape their attention either.

B. The Decision in Teoh

Both McHugh J’s dissent in the original Teoh67 decision, together with Gummow J’s obvious concern with the doctrine, expressed at the Lam hearing,68 gave some insight into their desire to revisit the principles upon which Teoh was decided. However, Lam was not actually a case on point which enabled them to effectively re-consider the judgment itself, as it contained no legitimate expectation arising from the act of treaty ratification. The decision itself was decided in a post-Teoh context,69 and consequently consideration of the relevant Convention had already been built into the decision-making process. Not only did the applicant not rely upon Teoh in that sense, but the respondent actually rejected that it had been relied upon also. Nonetheless, McHugh and Gummow JJ employed a strained logic to justify their consideration of the doctrine in Teoh:

---

66 Id at 302–303; (2003) 195 ALR 502 at [80].
67 Teoh, above n1.
68 Re MIMA; Ex parte Lam B33/2001 (24 June 2002) High Court of Australia Transcripts.
69 At the hearing, Mr Walker QC, Counsel for the Applicant referred to it as being a decision made ‘ante-Teoh’: Re MIMA; Ex parte Lam B33/2001 (24 June 2002) High Court of Australia Transcripts, at 5.
Counsel for the Minister disclaimed any direct attack on *Teoh* because, as he understood it, the applicant did not rely upon that case. … Nevertheless, the applicant’s submissions invited comparison with what was decided in *Teoh*. In particular, the applicant relied upon a strand in the reasoning of Mason CJ and Deane J, and Toohey J. This was that the absence of any particular “psychological effect” on the affected person or any other individual of knowledge of Australia’s adherence to the international obligations under the treaty in question was no impediment to the decision that there was a want of procedural fairness [footnote excluded].

The actual reliance on the point established in *Teoh* regarding the absence of an actual subjective awareness or knowledge of the legitimate expectation was actually not relevant to the facts in *Lam*. The expectation was an actual expectation implied through the correspondence sent to Mr Lam by the relevant departmental officers. Notwithstanding this fact, references to “psychological effect” at the earlier hearing were taken as a sufficient basis for their Honours to venture over the state of mind element considered in *Teoh*. In doing so, they sought to distinguish the expectation in *Teoh* from other ‘legitimate’ or ‘reasonable’ expectations, which included expectations involving ‘an actual or conscious appreciation that a benefit or privilege is to be conferred or a particular state of affairs will continue’. Their Honours considered that:

> It is one thing for a court in an application for judicial review to form a view as to the expectations of Australians presenting themselves at the gates of football grounds and racecourses. It is quite another to take ratification of any convention as a “positive statement” made “to the Australian people” that the executive government will act in accordance with the convention and to treat the question of the extent to which such matters impinge upon the popular consciousness as beside the point.

Even more pointedly, their Honours made the following claim:

> *Haoucher* does not stand beside *Teoh*. In the former case there was a statement made in the Parliament bearing immediately upon the exercise of the particular power in question. In *Teoh* there were in the Convention various general statements and there was no expression of intention by the executive government that they be given effect in the exercise of any powers conferred by the Act.

This point is directed to an issue of controversy that arose in the aftermath of *Teoh*, namely, whether *Teoh* involved the application of, or a significant extension of, the decision in *Haoucher*. This issue draws on the categorisation of treaty ratification

---

70 (2003) 195 ALR 502 at [84]–[85].
71 Id at [91].
72 Id at [95].
73 Id at [96].
with a considered statement of government policy. In *Haoucher*, a legitimate expectation was held to have existed on the basis of the contents of a published policy statement. Commentators such as Allars and Twomey have argued that *Teoh* is consistent with the earlier decision in *Haoucher*, on the basis that the act of ratification is akin to a considered statement of policy. Other commentators, such as Aronson and Dyer, have argued to the contrary. It would appear that the present High Court have serious reservations also in relation to that analogy.

Essentially the concern with *Teoh* rested on the fact that a significant expectation could arise by virtue of the act of ratification alone, and even in cases where an individual affected by a decision had no knowledge of the expectation itself. Their Honours did, however, accept the more conventional effects of an unincorporated treaty in domestic law, including its role in assisting with statutory interpretation, and where ratification is coupled with an additional step in the conduct of foreign affairs. What ultimately emerges as their major concern with the decision, is the fact that they interpret it as introducing mandatory considerations relevant to procedural fairness, and consequently, as overstepping the legitimate bounds of judicial review derived from the Constitution:

> [I]n the case law a line has been drawn which limits the normative effect of what are unenacted international obligations upon discretionary decision-making under powers conferred by statute and without specification of those obligations. The judgments in *Teoh* accepted the established doctrine that such obligations are not mandatory relevant considerations attracting judicial review for jurisdictional error. The curiosity is that, nevertheless, such matters are to be treated, if *Teoh* be taken as establishing any general proposition in this area, as mandatory relevant considerations for that species of judicial review concerned with procedural fairness.

The reasoning which as a matter of principle would sustain such an erratic application of “invocation” doctrine remains for analysis and decision. Basic questions of the interaction between the three branches of government are involved. One consideration is that, under the Constitution (s61), the task of the Executive is to execute and maintain statute law which confers discretionary powers upon the Executive. It is not for the judicial branch to add or vary the content of those powers by taking a particular view of the conduct by the Executive of external affairs. Rather, it is for the judicial branch to declare and enforce the limits of the power conferred by statute upon administrative decision-makers, but not, by reference to the conduct of external affairs, to supplement the criteria for the exercise of that power. [footnote excluded, emphasis added.]

With respect, the first italicised statement in the above quote contains an incorrect interpretation of the judgment in *Teoh*. Nowhere in that decision did any of the majority judges consider a ratified convention as imposing a mandatory

---

75 See Allars, above n5 at 224–225; Aronson & Dyer, above n42 at 327–328.
77 Allars, above n5 at 224-225.
78 Twomey, above n5 at 353–354.
79 Aronson & Dyer, above n42 at 328.
80 Id at [100].
81 Id at [101]–[102].
consideration upon decision-makers — merely that any decision inconsistent with the convention’s terms would necessitate the need for a hearing by the individual concerned. Indeed, in this respect, it appears that their Honours have adopted a rather strained reading of the decision itself. However, their concern is obviously with respect to the practical effect of the Teoh principle, which makes a convention a relevant consideration that must be taken into account by a decision-maker. The distinction that needs to be made (and which is not made by their Honours) is that the real effect of Teoh is the creation of a procedural obligation where departure from a treaty is intended. In that sense, the treaty becomes a mandatory relevant consideration. That, however, is quite distinct from asserting that the treaty itself constitutes a mandatory consideration that requires substantive compliance with its terms.

The procedural obligation in Teoh is activated upon a decision to act inconsistently with a treaty, and is actually premised on the freedom of a decision-maker to act in non-compliance with the treaty’s terms. It may be that a minimal (in the sense of being procedural) negative guarantee for the substantive terms of a treaty is effected by the decision in Teoh. However, this is entirely consistent with the Court’s traditional approach to the relevance of international law in Australia. Principles, that include the presumption that parliament does not intend to act in contravention of its international obligations, and that parliament must be expressly clear when intending to remove common law rights and freedoms, have long been used by Australian courts in relation to the interpretation of statutes. While Teoh concerned the actions of the executive, rather than parliament, a parallel can be drawn between the minimal, procedural guarantee identified in Teoh and the interpretive rules mentioned.

With respect to the second italicised statement above, the focus is misleadingly placed on the discretionary power as having been modified, rather than on the modification to the content of procedural fairness (which concerns the exercise of the power, rather than its actual scope). These matters are distinct, though it appears that McHugh and Gummow JJ have treated procedural fairness requirements based on ‘external affairs’ as involving an impermissible limitation being placed on the exercise of discretionary power. Given the acceptance by their Honours in Plaintiff S157/2002 v Commonwealth82 that breaches of procedural fairness constitute jurisdictional error, it appears that the element of ‘external affairs’ is what troubled their Honours in Lam. Indeed, the second italicised statement referred to is extremely difficult to reconcile with several lines of authority regarding the use of international law, as well as the court’s role in identifying the scope of statutory power.

C. Mr Lam’s Application

Having taken 38 paragraphs to consider issues not directly raised in argument, McHugh and Gummow JJ took only four paragraphs to address the application of Mr Lam. In doing so, they accepted that the failure to carry through with an act that

82 (2003) 195 ALR 24 at 37 (Gaudron, McHugh, Gummow, Kirby and Hayne JJ), citing Re Refugee Review Tribunal; Ex parte Aala (2000) 204 CLR 82.
the departmental officer had said would be done, involved the failure to observe an expectation reasonably attributable to the applicant.\(^{83}\) However, rather than choosing not to grant a remedy which would have been available given the discretionary nature of the remedies sought, their Honours refused to find that a denial of natural justice had occurred, dismissing the application. In this respect, their Honours came close in approach to that of Gleeson CJ. Like the Chief Justice’s, their reasoning is also subject to the criticism that their approach incorporates a notion of detriment and unfairness in outcome, despite their acceptance of fairness in procedure as being the fundamental issue for the Court.\(^{84}\) They addressed the issue of a denial of natural justice as follows:

> But the failure to meet that expectation does not reasonably found a case of denial of natural justice. The notion of legitimate expectation serves only to focus attention on the content of the requirement of natural justice in this particular case. The ends sought to be attained by the requirement of natural justice may be variously identified. But at least in a case such as this the concern is with the fairness of the procedure adopted rather than the fairness of the outcome. It is the decision-making process not the decision, as Lord Brightman put it. What is delivered by the requirement of natural justice is the right to a hearing, a technical expression in law, before action is taken.

The applicant by the statement in the letter to him … did not acquire any vested right to oblige the Department to act as it indicated … It was not suggested that in reliance upon that letter the applicant had failed to put to the Department any material he otherwise would have urged upon it. Nor was it suggested that, if contacted, the carers would have supplemented to any significant degree what had been put already in the letter of 17 October 2000. The submission that the applicant, before the making by the Minister of his decision, should have been told that the carers were not to be contacted, thus lacks any probative force for a conclusion that the procedures so miscarried as to occasion a denial of natural justice [emphasis added, footnote excluded].\(^{85}\)

Essentially this approach rests upon notions of reliance and detriment, which go extremely close to entailing an examination of the merits of a case. It appears that this has probably resulted from the fact that the issue of breach has been conflated with the issue of discretion to grant the remedy sought. The more appropriate manner of considering the issues before the Court would have been to consider the issue of breach separately from the issue of remedies, as the latter question requires the Court to consider substantive issues to a degree. The discretionary nature of the remedies available lends itself to such an approach, and has the added benefit of avoiding a situation where principles more apt to a private law context are not employed in a public law setting where judicial review is constitutionally limited. Had McHugh and Gummow JJ been serious about their concern for fairness within the process rather than the outcome, they would have focussed on the absence of an opportunity to make submissions or present new material, rather than on the absence of evidence of any additional material itself that may have been presented.

\(^{83}\) Id at [103]–[104].  
\(^{84}\) Id at [105].  
\(^{85}\) Id at [105]–[106].
The former is concerned with the procedural aspects of the decision-making process, the latter with the substance or merits of the decision. Indeed, counsel for the applicant mentioned in argument potential matters that new material could have been put forward. In particular, evidence of the language background of the Lam children and of their ability to effectively integrate into Vietnamese society were matters that could have been further considered.86

What is meant, however, by the statement that reliance, detriment or unfairness in outcome are more appropriately considered in relation to the discretionary remedy? This notion draws upon the authority of previous High Court decisions in *Re Refugee Tribunal; Ex parte Aala*,87 and *Stead v State Government Insurance Commission*.88 Each of these cases is relevant to the discretionary nature of the remedies available for denials of natural justice, with the decision in *Aala* specifically concerned with the constitutional writs available under s 75(v) of the Constitution. Interestingly, the case of *Stead* is of only limited relevance in an administrative law context as it concerned the discretion of a superior court to order a new trial where a party had been denied an opportunity to make submissions at a previous criminal trial. To the extent that the unanimous judgment concerned the issue of discretionary remedies for breaches of natural justice, however, the High Court has affirmed its relevance in *Aala*.89 The main principle taken from *Stead* is as follows:

> [N]ot every departure from the rules of natural justice at a trial will entitle the aggrieved party to a new trial. By way of illustration, if all that happened at a trial was that a party was denied the opportunity of making submissions on a question of law, when, in the opinion of the appellate court, the question of law must clearly be answered unfavourably to the aggrieved party, it would be futile to order a new trial.

Where, however, the denial of natural justice affects the entitlement of a party to make submissions on an issue of fact, especially when the issue is whether the evidence of a particular witness should be accepted, it is more difficult for a court of appeal to conclude that compliance with the requirements of natural justice could have made no difference.90

This case obviously did not involve the exercise of judicial review of administrative action, as in *Aala*. In that more recent case, Kirby J articulated the principle in that context as follows:

> Once the applicable breach is proved, the victim of the breach is ordinarily entitled to relief. It is only where an affirmative conclusion is reached, that compliance with the requirements of procedural fairness “could have made no difference” to the result, that relief will be withheld. This Court has emphasised that such an outcome will be a rarity. It will be “no easy task” to convince a court to adopt it.91

---

86 See, *Re MIMA; Ex parte Lam* B33/2001 (24 June 2002) High Court of Australia Transcripts.
87 (2000) 204 CLR 82.
88 (1986) 161 CLR 141.
89 (2000) 204 CLR 82 at [4].
90 (1986) 161 CLR 141 at 145.
In considering the rationale underlying the principle, Kirby J continued:

> It does not assure the victim of the breach of ultimate success. But it does assure that person of the privilege belonging to all those affected by the deployment of power by officers of the Commonwealth. This is that such officers will only act in accordance with their lawful mandate. The exception, accepted by Stead, is held in reserve to guard against insignificant, purely formal and immaterial mistakes. Unless the breach can be so classified, the person affected who claims the writ is normally entitled to relief. [Emphasis added.]

If one applies this principle to the facts of Lam, one may acknowledge the technical breach of procedural fairness at the same time as acknowledging a principled basis for denying the remedy as sought by Mr Lam. While the applicant could have made further submissions, in light of the substantial submissions already presented on his behalf, it is unlikely that any further evidence would have modified the outcome of the decision-making process. In this sense, one may distinguish the facts in Lam from those in Aala, where it could not be so concluded. The problem with the decision in Lam, is that it appears to be distinguishable from Aala on the basis that no reliance or detriment could be found to exist on the facts in Lam. However, the absence of detrimental reliance in Lam is only properly seen as being relevant to whether or not the remedy sought should indeed be granted.

To view the issue in this light is to acknowledge that unfair procedures need not necessarily result in an unfair outcome. However, the Court in Lam effectively adopted an approach that unfair procedures were to be treated as fair provided that a substantively unfair result did not occur. Such retrospective reasoning should be questioned, especially given that it may encourage poor administrative practices. Indeed, the Court’s role in identifying and developing principles in this area should be premised on notions of good administration that recognise the ‘public’ nature of administrative law. Administrators should be able to look to general rules, guidelines and benchmarks that offer the minimum standard with which decision-makers must comply in order to ensure consistency and fairness in the circumstances of each case. Principles that rest on the subjective assessment of the decision-maker as to whether unfairness may result (as in Lam), or upon a retrospective assessment of unfairness by the courts, do little to encourage such practices in administrative decision-making.

### 6. Justice Hayne

The judgment of Justice Hayne is a particularly interesting one. In making reference to the criticisms or concerns of McHugh and Gummow JJ, the judgment is unhelpful, and merely adds further confusion to matters that were not actually raised by the case. However, Hayne J is more direct than any of the other judges in

---

91 Id at 130–131.
92 Id at [131].
93 On this point, see the discussion that took place during the hearing: Re MIMA; Ex parte Lam B33/2001 (24 June 2002) High Court of Australia Transcripts.
addressing some of the underlying concerns, and consequently, in pointing to the basis upon which many of the concerns were raised. Hayne J indicates very early in his judgment that he is not in favour of the development of the doctrine of legitimate expectation over time, from one linked to interests attracting procedural fairness, to one that directly affects the content of procedural fairness. Consequently, his approach limits the relevance of the legitimate expectation. As his Honour reasons:

If the procedure was fair, reference to expectations, legitimate or not, is unhelpful, even distracting.

Confining the description of the events and circumstances to those I have mentioned omits reference to two critical matters. First, the applicant accepted that, but for what the Department said it would do, procedural fairness would not have required the Department to interview the carer. Secondly, he did not suggest that, had he known that the Department would not contact his children’s carer, he would have submitted any additional material or argument.94

Taking this approach, Hayne J reduced the content of procedural fairness by removing the relevance of the legitimate expectation in determining what fairness required. The fact that Mr Lam had been given a sufficient opportunity to be heard satisfied the requirements of procedural fairness, and Mr Lam’s conduct was not modified by any failure to follow the procedure of contacting the children’s carer:

The applicant was given the opportunity to submit, and did submit, all the material and all the arguments that he wanted to submit before the decision was made. The Department’s statement that it intended to contact the children’s carer did not cause the applicant to alter his conduct in any way. In particular, it did not cause him to refrain from substituting for, or adding to, the material and argument he had already submitted.95

In this sense, Hayne J is able to avoid some of the criticism levelled against the preceding judgments on the basis that he does not rely on notions of reliance and detriment to avoid the procedural requirements that arise by virtue of the legitimate expectation. The expectation itself is irrelevant in determining the content of procedural fairness. However, this approach is almost artificial, as an undertaking made by a decision-maker, whether or not it is seen as generating a ‘legitimate expectation’, must, like a policy statement or other conduct, affect the content of procedural fairness. The legitimate expectation is merely a label attached to the action or conduct of a decision-maker relevant to the decision-making process. Its utility, even necessity, as a label in this context is certainly open to question, but one cannot deny the relevance of the action or conduct in determining the requirements of procedural fairness by simply rejecting the role of legitimate expectation here.

When the approach of Hayne J is viewed in this light, the same criticisms levelled at the other judgments may also be levelled at his Honour’s. The notion of

---

94 (2003) 195 ALR 502 at [111]–[112].
95 Id at [114].
what procedural fairness requires is assessed retrospectively, and the actual presence or absence of some indication that an individual would have acted differently is the ultimate issue. In other words, the presence of actual detriment becomes significant. This approach comes extremely close to review on the merits. Judicial review for want of procedural fairness should be limited to the procedural aspects, rather than substantive questions. What should really have been at issue in Lam, was not whether there existed other evidence which Mr Lam would have presented, or additional steps he would have taken in light of the failure to follow a stated procedure, but whether he was denied an opportunity to do so. The Court’s role was not to review the quality or merits of his case but the fairness of the procedures taken.

Despite having taken a narrow view of legitimate expectation, and thereby limiting its relevance in determining the content of procedural fairness, Hayne J qualifies that statement, and considers what would be the outcome were the letter to Mr Lam capable of generating a legitimate expectation:

[It is enough to say that … departure from it, where it is accepted that neither the expectation nor departure from it affected the course which the applicant pursued, gives no ground for relief. He was afforded a full opportunity to be heard. The Department’s letter raised no new matter to be taken into account in making the impugned decision. … Unlike Teoh, this was not a case where the course of decision-making could be said to have diverged from any announced policy to be taken into account in making the relevant decision.]96

A. The Decision in Teoh

Unlike McHugh and Gummow JJ, Hayne J did not go into a detailed critique of the Teoh decision. He did, however, indicate that there were many issues relating to legitimate expectations generally that remained to be considered, and that the application of the Teoh principle stood to be limited or confined in some way. However, Hayne J was not as quick to reject the reasoning in Teoh altogether, but left open the potential for its review, and potential refinement and limitation. On these points, his Honour made the following statements:

[Used in its broader sense] legitimate expectation is a phrase which poses more questions than it answers. What is meant by “legitimate”? Is “expectation” a reference to some subjective state of mind or to a legally required standard of behaviour? If it is a reference to a state of mind, whose state of mind is relevant? How is it established? These are questions that invite close attention to what is meant by legitimate expectation and what exactly is its doctrinal purpose or basis. Not all are dealt with explicitly in Teoh. At the least they are questions which invite attention to the more fundamental question, posed by McHugh J in Teoh, of whether legitimate expectation still has a useful role to play in this field of discourse now that it has served its purpose in identifying those to whom procedural fairness must be given as including more than persons whose rights are affected.

It may be that, for the reasons given by McHugh and Gummow JJ in this matter, Teoh cannot stand with the Court’s earlier decision in Haoucher v Minister for

96 Id at [122].
Immigration and Ethnic Affairs. It may also be that further consideration may have to be given to what was said in *Teoh* about the consequences which follow for domestic administrative decision-making from the ratification (but not enactment) of an international instrument. [Footnotes excluded.]

Though Hayne J raised the presence of several matters still warranting further consideration, or perhaps more accurately, matters that the Court is eager to reconsider, he accepted that *Lam* was not the case in which to do so. Accordingly, he refrained from offering pages of obiter comments and accepted that such matters ‘need not be answered in this case’.

7. **Justice Callinan**

Like McHugh and Gummow JJ, Callinan J spent only four paragraphs addressing the primary issue before the Court. Like each of the other judges, he dismissed the application on the basis that Mr Lam could not point to any additional material that he would have put to the decision-maker had he known that the children’s carer had not been consulted. As his Honour stated:

> In my opinion, what is fatal to the applicant’s claim here is that he was unable to demonstrate that there was any material that he could have put before the respondent which was either not already in the respondent’s hands, or which might have influenced the respondent to decide his case differently. That he might have liked to have had a further opportunity to repeat what he had already said, or to advance the same argument differently or more emphatically is not to the point and cannot avail him.

Thus, though it is only implicit in the judgment of Callinan J, the notion of detriment as a requirement of a denial of procedural fairness, is present here also. Consequently, this judgment is also subject to the same criticism as each of the other separate opinions on this point.

A. **The Decision in Teoh**

Justice Callinan devoted 12 paragraphs of his judgment to blatant criticism of the decision in *Teoh*, premised on the fact that he was not convinced by counsel for the applicant, that the applicant did not need to rely on the decision. Callinan J began this attack with the following statements:

> In my opinion, the expression “legitimate expectation” is an unfortunate one, and apt to mislead. In the case of *Teoh*, it was, with respect, a complete misnomer. … Moreover, the necessity for the invention of the doctrine is questionable. The law of natural justice has evolved without the need for recourse to any fiction of “legitimate expectation”.

97 Id at [121]–[122].
98 Id at [122].
99 Id at [149].
100 Id at [140].
While Callinan J clearly had concerns with the underlying reasoning of the majority in *Teoh*, his major criticism was directed against the ‘objective assessment’ of an expectation espoused most strongly in that case in the judgment of Toohey J. His rejection of that approach was argued as follows:

I would observe with respect, that an “undertaking” presupposes a recipient of it, just as an “engendering” will be meaningless unless it has an effect upon the mind of someone.

It seems to me, with respect, that if a doctrine of “legitimate expectation” is to remain part of Australian law, it would be better if it were applied only in cases in which there is an actual expectation, or that at the very least a reasonable inference is available that had a party turned his or her mind consciously to the matter in circumstances only in which that person was likely to have done so, he or she would reasonably have believed and expected that certain procedures would be followed.\(^{101}\)

To give further indication of the level of concern (one may say obsession), with issues not even directly argued before the court in *Lam*, Callinan J made the extraordinary statement that he could not help but make further reservations regarding *Teoh*, ‘before moving to the facts of this case’. Those further reservations related to the relevance of unincorporated treaties in *Teoh*, and on this point Callinan J’s view was stated as follows:

The fact remains that the Convention is not part of Australian law…. In consequence, the view is open that for the Court to give the effect to the Convention that it did, was to elevate the executive above the parliament. This in my opinion is the important question rather than whether the executive act of ratification is, or is not to be described as platitudinous or ineffectual. [Footnotes excluded.]\(^{102}\)

Before actually moving to the facts of the case, Callinan J also confirmed his support for the views expressed by McHugh and Gummow JJ regarding the absence of substantive rights in relation to legitimate expectations.

### 8. Conclusion

Commentators should not resist the call to further refine and articulate the notion of legitimate expectations within administrative law. For many years, several commentators have questioned its utility beyond the identification of an interest sufficient to trigger the application of procedural fairness. The extent to which it plays an important role in helping to determine the content of procedural fairness warrants further attention. It may well be that the concept of a legitimate expectation only confuses the matter. Perhaps the more appropriate issue is the action or conduct of the decision-maker and how that in itself should be viewed as affecting the content of procedural fairness. In this respect, an undertaking or representation, an action or policy, may each shape what fairness requires in the

---

101 Id at [144]-[145].
102 Id at [147].
circumstances. The presence of a legitimate expectation on the part of an individual concerned may be entirely irrelevant, representing an artificial construct that adds little or nothing to the ultimate question. In this respect, the personal or subjective knowledge or awareness of the policy, conduct, act or undertaking may be of no relevance at all, or perhaps only of secondary importance. Such an approach more appropriately reflects the particular considerations relevant in public law — the development of principles based on notions of good administration, and an emphasis on the conduct of government decision-makers.

However, this issue deals with the broader role of legitimate expectations generally. The Teoh doctrine is a separate, but related issue. The extent to which the previous minority in Teoh may have come to represent a majority of the present High Court, means that the decision is likely to be severely overhauled if the Court is given the opportunity to reconsider it. Whether it is modified substantially, or limited in the sense outlined briefly by Hayne J, remains to be seen. What is absent from the discussion in Lam, however, is any consideration of the relevance of other factors (in addition to the act of ratification) to the generation of a legitimate expectation. Indeed the relevance of the procedures adopted since 1996 regarding treaty-making, as well as the scheduling of specific conventions to HREOC and other Acts, was not addressed. Even if the Court were to circumvent Teoh, it would be difficult for the Court to deny that legitimate expectations arise in cases where Parliament has endorsed the ratification in some direct way. Consequently, Teoh may prove to be more resilient than at first anticipated.

Finally, what is perhaps of most concern is the use of private law concepts in the public law context. In particular, the use of notions of reliance and detriment in procedural fairness cases is a disturbing trend. It brings the Court far too close, even into, a review of the substantive merits of a decision, and moves beyond the legitimate bounds of judicial review. What the Court should have done in Lam was to have acknowledged the denial of procedural fairness, but chosen not to grant the remedy sought. This approach acknowledges the absence of fair procedures, putting the decision-maker on notice, but refrains from quashing the decision in cases where the breach was only minor or technical. As an approach, it relies heavily on the discretionary nature of the remedies available in administrative law, which Gleeson CJ himself referred to in Lam. The Court is right to clarify and further refine the notion of legitimate expectations, but in doing so it should be mindful of its rightful role in the conduct of judicial review of administrative action.

103 See Lacey, above n13 at 233–238.
Call for Papers for the 50th Anniversary Commemorative Issue of the Sydney Law Review

Legal Education & Legal Scholarship: Canon, Critique or Crisis

“[T]he educational institution, which plays a crucial role in the reproduction of the distribution of cultural capital and thus in the reproduction of the structure of social space, has become a central stake in the struggle for the monopoly on dominant positions.”

PIERRE BOURDIEU (1989)

“Conventional legal education puts blinders on the students, enabling them to tread doggedly a well-trodden path of professional success, and generates forms of scholarship that accept the borders of the path as the boundaries of the legal universe.”

RICHARD POSNER (1993)

“To date, not much has been required of legal progressives who supported, or at least, went along with minority activists…a few white women hired, scholars of colour at least interviewed, the insertion of a couple of readings on racism and feminism in conventional core materials, the approval of one or two elective courses…[but] [f]eel-goodism is not a politics, let alone a politics of change”.

SHEILA MCINTYRE (1993)

“[T]alk of a crisis in scholarly legal writing is fundamentally misconceived…the state of legal writing in Australia promises [the Sydney Law Review] an even greater future.”

MICHAEL TILBURY (2003)

To celebrate the Sydney Law Review’s 50th Anniversary, the Review is pleased to invite scholars in all disciplines and of all nationalities to submit papers for a Special Commemorative Issue of the Review to be published in December 2004. Submissions are sought that engage with the theme of this Commemorative Issue: legal education and legal scholarship. In particular, the Editorial Board would welcome writing on one or more of the following topics:

• the role of legal scholarship;
• historical or contemporary trends in legal research;
• current issues in legal education; and/or
• the relationship between legal education and research.

The Sydney Law Review is a refereed, peer reviewed journal with a wide readership both in Australia and internationally. It is published quarterly by Lawbook Co. and The University of Sydney Law School. The Review is edited by an Editorial Board of academic staff and a Student Editorial Committee.

The Editorial Board will consider all papers received by 1st August, 2004 for publication in December 2004. All inquiries should be directed to Ms Joanna Howse, Sydney Law Review Coordinator: Phone: + 61 2 9351 0284; Fax: + 61 2 9351 0407; Email: joannah@law.usyd.edu.au. For more information, including submission and style guidelines, please visit the Review’s web site at http://www.law.usyd.edu.au/~slr.