## Contents

Foreword ................................................................. vii
From the Editors ......................................................... ix
List of Subeditors ....................................................... xi
List of Authors .......................................................... xiii
List of Reviewers ....................................................... xxi
Cover Art ................................................................. xxvii

The Breaking of the ‘Great Australian Silence’: How and Why the Writing of Indigenous Australian History has Changed over the Last 40 Years ....................................................... 1
   Caroline Beasley

Heritability Estimates of Complex Intelligence and Associated Genetics ................................................. 15
   Nyomi Bourbous

Tradition and the Art of Modern India ................................. 25
   Kieran Browne

The Promise and Failure of King Amanullah’s Modernisation Program in Afghanistan ........................................... 35
   Andrew Chua

Mechanistic Explanation: Some Limits and their Significance ................................................................. 51
   Dominie Dessaix

Dependent Yet Defiant: The Implications of Unilateralism in Iraqi Kurdistan ................................................... 61
   Sebastian Klich

Classification of Arid & Semi-Arid Areas: A Case Study in Western Australia ........................................ 79
   Guy Leech

Framing Gender Inequality: Millennium Development Goal 3 and the Post-2015 Agenda ............................................ 101
   William Lutwyche

Domestic and Sexual Violence Related Asylum Claims in the Refugee Review Tribunal ....................................................... 115
   Amanda Joyce Neilson
Susan Yuan Wen Nheu

A Rock Amidst Turbulent Waters: China’s Resilient Centre and its Machiavellian Management of Protests and Petitions... 145
Sam Osborne

Was there a Political Dimension to the Irish Rockite Movement of 1821 to 1824? .................................................. 155
Rebecca Preston

Mainstreaming Restorative Justice in South Australia’s Criminal Justice System: A Response to the Over-Representation of Indigenous Offenders .................................................. 169
Alexandra Smith

Officer-Facilitated Homicide: Assessing the Deaths of Mentally Ill Offenders as a Result of Police Practices ............ 187
Benjamin A Smith

Persistent Levels of Poverty in India: Urgent Action Needed to End Suffering ......................................................... 201
Harita Sridhar

Gender-based Differences in Complimenting Behaviour: A Critical Literature Review ............................................ 213
Nan Sun

Women’s Resistance Efforts in Nazi Germany 1939–45: HerStory ................................................................. 223
Jillian Wales

Comparing the Energy Consumption of Logistics Chains in Traditional Retailing and E-commerce for Popular Products in the Context of China ......................................................... 243
Wenjia Wang

Sanitary Paper Consumption: A Material Stocks and Flows Analysis at Fenner Hall, 2012 ........................................... 269
Huan Zheng, Yale Wong
Foreword

Welcome to the 2013 edition of the *ANU Undergraduate Research Journal (AURJ)*.

It is a pleasure to see our undergraduate students publish with such depth and breadth. I am always impressed by the range of disciplines represented in this outstanding journal.

In the following pages you will read about how the online movement has profoundly influenced the logistics and energy consumption of shopping, Nyomi Bourbous looks into the heritability of intelligence, and Jillian Wales presents some fascinating case studies of women's dissidence in Nazi Germany.

Other topics include the justice system, modernisation in Afghanistan, poverty in India, and even a detailed study of toilet paper usage by the residents of Fenner Hall.

I hope you will delight, as I do, in the quality and real-world relevance of the research presented here.

Professor Ian Young AO
Vice-Chancellor and President
From the Editors

We were both impressed and somewhat dismayed as we sifted our way through the various submissions to the ANU Undergraduate Research Journal (AURJ) 2013. Impressed at the quality and depth of the research being conducted by our undergraduate peers; dismayed that we had to turn away so many quality manuscripts.

We invite readers to discover the breadth of the research being undertaken at the ANU, exhibited herein in the diversity and range of subjects explored in this volume. The multifaceted nature of the ANU research is clearly illustrated, with offerings from international politics, music, historical resistance movements, the environmental efficiency of online shopping and, closer to home, the use of sanitary paper in one of the University’s residential halls.

We congratulate the students for their excellent work, and their commitment and tenacity as they proceeded through the review process. We trust that they have been provided with valuable experience for their future career paths, wherever they may lead.

We thank the numerous reviewers who gave their time to offer comments and constructive criticism to the authors, ensuring that the end results would be of the highest possible quality. Particular thanks go to the reviewers that offered to mentor the students through the process of revising their manuscripts.

Thanks also to our excellent team of subeditors, whose work has ensured the smooth running of the review process to a tight timeframe. We also thank the Vice-Chancellor, Professor Ian Young, and Professor Richard Baker, the Pro Vice-Chancellor Student Experience for their support for AURJ. Thanks to Stefan Nesteroff, Project Coordinator, for his guidance and encouragement, and to ANU Press for their advice and support.

We have enjoyed collating these pages and the opportunity this afforded us to be exposed to ideas and issues beyond our own niches. We hope that readers will also enjoy this foray into work that reflects and bolsters ANU’s reputation for high calibre research.

Kazi Rahman, Kaveenda Samarasinghe, Sarah Tynan

AURJ 2013 Editorial Team
List of Subeditors

Shea Andrews
Daniel Wei Boon Chua
Christina Delay
Emily Duthie
Amanda Edworthy
Sara Perley
Isabel Roper
Tennille Sibbritt
Dinar Thalieb
Anna Tsalapatanis
List of Authors

Caroline Beasley

Caroline moved to Canberra from Tasmania to study Arts/Law in 2011. Her interest in Australian history led her to explore the way in which Indigenous Australians have so often been excluded from the Australian story. She is also interested in the status of Indigenous peoples in Australia from a legal perspective, and is looking forward to continuing her involvement with the ANU Law School’s Ready 4 Recognition group after coming back from exchange at Trinity College Dublin in July 2014. She hopes to pursue an honours project in this area in the future.

Email: u4855615@anu.edu.au

Nyomi Bourbous

Science in general has always intrigued Nyomi, but when she first came to the ANU she had no idea of the field in which she wanted to specialise. In her first and second years Nyomi was introduced to genetics and was hooked, believing that she had found her calling. In addition, she is also interested in evolutionary theory specifically pertaining to humans. She is currently in her final undergraduate year and has aspirations to complete honours in biological anthropology next year. She adores studying and acquiring new knowledge and it is her aspiration that one day she may be able to carry out her own research.

Email: u4843355@anu.edu.au

Kieran Browne

Kieran is in his final semester of a Bachelor of Digital Arts. His research has been concerned with the intersection of art and science and the globalisation of modern art theory in the latter part of the twentieth century. He is interested in historical and non-Western art theories, especially that of Indian and ‘oriental’ modernisms that have been ‘discovered’ by Western art discourse in recent years. His art practice and development of a theoretical framework owes much to the expansion of art discourse. Kieran hopes to continue working as a practising artist after his university studies.

Email: kierantbrowne@gmail.com
Andrew Chua

Andrew graduated from the ANU with a Bachelor of Arts in July 2013, majoring in political science and ancient history. Andrew will be starting Bachelor of Arts (Honours) program from February 2014, with his thesis based on the politics of state-building in the Occupied Territories.

Email: u4808729@alumni.anu.edu.au

Dominie Dessaix

Dominie is a third-year Bachelor of Arts student, majoring in linguistics and philosophy. Her main interests are philosophy of mind and psychology, cognitive science, philosophy of science, and philosophy of language. She is fascinated by questions involving interdisciplinary work, or that have answers bearing on work done in many disciplines. Some of these questions are the nature of explanation, consciousness, and the role of language in what makes human activity unique. Dominie is also interested in how findings from linguistics might inform, and perhaps ‘naturalise’, the philosophy of language, and how the biological sciences might interact with linguistics, philosophy, and the humanities more generally. She hopes to begin honours mid-semester 2014, either single honours in philosophy, or combined honours in philosophy and linguistics.

Email: u5024827@anu.edu.au

Sebastian Klich

Sebastian is an international relations honours student specialising in Middle Eastern studies. His thesis uses a constructivist framework to analyse the identity and contested sovereignty of the Kurdistan region in northern Iraq, hence he chose to focus on the strategic and regional implications of the Kurdistan Regional Government’s unilateral natural resource management for this essay. After completing honours he hopes to complete a PhD at the ANU, conducting research into how social theory in international relations can provide a better understanding of how unrecognised states function in international society.

Email: sebastian.klich@anu.edu.au
List of Authors

**Guy Leech**

Guy is an environmental science student (and tutor) who came to the ANU to study animals, and ended up studying climate. Despite dropping out of high school, he's enjoying himself now. His future plans include honours studying past climate change, and delicious cheese.

Email: guy.leech@anu.edu.au

**William Lutwyche**

After working in Western and Northern Africa in 2009, William moved to the ANU from Tasmania to pursue his interest and passion for International Development. Throughout his time in Canberra, William has worked extensively with the Youth-led development NGO the Oaktree Foundation and completed internships at the Australian Council for International Development and also a human rights based internship in Cape Town, South Africa. During 2013, William was awarded First Class Honours in Development Studies. His work in this journal formed part of the coursework required for this award. This Gender and Development based article links strongly with William's thesis topic but also is related to his experiences and knowledge of social development. Beyond 2013, he plans to search for jobs in development consultancy both in Australia and abroad. Additionally, he hopes to continue his passion for research to better understand and formulate solutions to some of the world’s most complex social problems.

Email: u4681197@anu.edu.au

**Amanda Joyce Neilson**

Amanda will graduate from a Bachelor of Arts/Laws (Hons) in July, 2014. She has recently completed her Honours Thesis in law, which focuses on the development of national security jurisprudence in the High Court of Australia. Amanda has been has involved in many aspects of student life during her time at the ANU. These have included leadership positions at Bruce Hall, ANU Students’ Association and ANU Student Media (Woroni). In 2014, Amanda will commence an internship in the civil law division of the Northern Australian Aboriginal Justice Association in Darwin, NT. She hopes to maintain her passion for social justice and community service throughout her career, wherever that may lead.

Email: amandajoycneilson@gmail.com
Susan Yuan Wen Nheu

Susan is currently in the second year of studying Bachelor of Music and Bachelor of Laws at ANU. Susan is studying classical singing and musicology as part of her music degree, and she is in the process of exploring different areas in law before deciding on a specialisation later on in her studies.


Email: susan.nheu@outlook.com

Sam Osborne

Sam Osborne is a sixth year combined Law/Asian studies (specialist) student currently completing his honours thesis in international law. His focus areas of study include Chinese studies and Mandarin, as well as international and human rights law. In 2012 he undertook a year of intensive language study at Tsinghua University, Beijing, and hopes to complete his Graduate Diploma in Legal Practice in 2014. His aim is to pursue a career in international law, either as a practising lawyer or in the field of policy advice relating to the development of adequate justice and governance frameworks in the Asia-Pacific region.

Email: samosbo@gmail.com

Rebecca Preston

Rebecca grew up in Sydney and moved to Canberra in 2010 to study a combined degree of Arts/Law. She is majoring in history and is completing her final subjects to finish her arts degree. She intends to do honours in history in 2014. Her first essay on the Irish Rockite movement was inspired by her exchange to Trinity College, Dublin, Ireland, which sparked her interest in early Irish rebellions and political history.

Email: u4839635@anu.edu.au
Alexandra Smith

Alexandra is in her fifth year of a double degree in Law and Development Studies. She is interested in protecting the legal rights of Indigenous offenders in the criminal justice system. She completed an Aurora Native Title Internship with the Aboriginal Justice Centre, Canberra, where she was exposed to the concept of restorative justice. During this internship she completed a proposal for the construction of a multijurisdictional Magistrate’s Court offering on-site social welfare services based on the principles of restorative justice. In the future, she hopes to work in criminal law on behalf of defendants.

Email: alexandra.smith.54@gmail.com

Benjamin A. Smith

Benjamin is an undergraduate student at the ANU College of Law, with an interest in criminal law and social justice. He has written a number of submissions to government agencies on issues relating to police powers, violent street crime, privacy, sexting, slavery and slaver-like conditions. These were undertaken in a private capacity with practicing academics and through Civil Liberties Australia. He has recently had papers accepted for publication in the Privacy Law Bulletin.

Email: u3302742@anu.edu.au

Harita Sridhar

Harita is a development studies/law student who wrote this policy brief as an assessment for a development course. She thoroughly enjoyed researching and writing the brief as the development of Global South nations is one of Harita’s passionate interests in which she hopes to pursue a career.

Email: harita.sridhar@anu.edu.au
Nan Sun

Nan graduated from the ANU with a B. Science/B. Arts double degree in December, 2013. Her submission to AURJ 2013 was initially written for a linguistics course she took at the ANU. Inspired by Lean In, a best seller written by Facebook's COO Sheryl Sandberg, Nan decided to approach her topic from a gender point of view. Nan plans to pursue an MBA degree in the US in the near future.

Email: nanapplesun@gmail.com

Jillian Wales

Jillian didn’t study history in her final years of secondary school, yet at the ANU it has become not only her arts major, but her passion. Over four years of study she has developed a particular interest in the diverse and complex history of World War II, principally in relation to the experiences of citizens living in Nazi Germany. German resistance, more specifically women’s resistance within Germany, is commonly downplayed as insignificant. Jillian seeks to dispel this notion by demonstrating that the efforts of dissident women should not be overlooked.

Email: jillianwales@hotmail.com

Wenjia Wang

Wenjia Wang is a student from the ANU College of Engineering & Computer Science, with a double major in mechanical and material systems, and renewable energy systems. She transferred to the ANU from Beijing Institute of Technology in July 2011. Her background in renewable energy stems from Wenjia's interest in studying the issues of resource shortages and sustainable development. Wenjia graduated in December 2013, and she is currently applying to postgraduate schools in the United States to further her study in renewable energy-related subjects.

Email: nicole.wangwj@gmail.com

Huan Zheng and Yale Wong

Huan Zheng is a third-year arts student majoring in environmental studies and sociology. She is passionate about biodiversity conservation, sustainability
and gender studies. Yale Wong is a student of geography and economics in the Bachelor of Philosophy (Honours) — Science. He has a passionate interest in sustainability issues, particularly relating to transportation planning and policy. As residents of Fenner Hall, Huan and Yale are acutely aware of the quantities of sanitary paper used, hence the material stocks and flows analysis conducted.

Email: u4928103@anu.edu.au and yale.wong@anu.edu.au
List of Reviewers

Ms Santhi Achuthan
The ANU College of Medicine, Biology & Environment
A/Prof. Bruce Baer Arnold
University of Canberra, School of Law
Dr Merridee Bailey
University of Adelaide, School of History & Politics
Dr Nida Bangash
The ANU College of Arts & Social Sciences
Dr Luisa Batalha
The ANU College of Medicine, Biology & Environment
Mr Adrian Bazbauers
The ANU College of Arts & Social Sciences
Ms Jennifer Beattie
The ANU College of Arts & Social Sciences
Dr Alison Behie
The ANU College of Arts & Social Sciences
Dr Chris Bishop
The ANU College of Arts & Social Sciences
A/Prof. Frank Bongiorno
The ANU College of Arts & Social Sciences
A/Prof. Wendy Bonython
University of Canberra, School of Law
Prof. Steve Dovers
*The ANU College of Medicine, Biology & Environment*

Prof. Steve Ellner
*The ANU College of Arts & Social Sciences*

Mr Nicholas Engerer
*The ANU College of Medicine, Biology & Environment*

A/Prof. Tony Foley
*The ANU College of Law*

Dr Karen Fox
*The ANU College of Arts & Social Sciences*

Dr Charlotte Galloway
*The ANU College of Arts & Social Sciences*

Ms Michal Glikson
*The ANU College of Arts & Social Sciences*

Dr Christian Goeschel
*The ANU College of Arts & Social Sciences*

Dr Jane Golley
*The ANU College of Asia & the Pacific*

Dr Patrick Guinness
*The ANU College of Arts & Social Sciences*

Dr Jill Guthrie
*National Centre for Indigenous Studies*

Ms Margaret Hutchison
*The ANU College of Arts & Social Sciences*
Mr Niamatullah Ibrahimi
*The ANU College of Asia & the Pacific*

Prof. Frank Jackson
*The ANU College of Arts & Social Sciences*

Dr Fiona Jenkins
*The ANU College of Arts & Social Sciences*

Ms Yimeng Jiang
*The ANU College of Engineering & Computer Science*

Dr Colin Klein
*The ANU College of Arts & Social Sciences*

Prof. Ian McAllister
*The ANU College of Arts & Social Sciences*

Ms Hannah McCann
*The ANU College of Arts & Social Sciences*

Dr Olivia Meehan
*The ANU College of Arts & Social Sciences*

Mr Brendan Moloney
*The ANU College of Engineering & Computer Science*

A/Prof. Barry Newell
*The ANU College of Engineering & Computer Science*

Prof. Daniel Nolan
*The ANU College of Arts & Social Sciences*

Dr Maria Nugent
*The ANU College of Arts & Social Sciences*
List of Reviewers

A/Prof. Molly O'Brien

*The ANU College of Law*

Ms Kate Ogg

*The ANU College of Law*

Ms Amanda Padovan

*The ANU College of Medicine, Biology & Environment*

Dr Gabriel Rabin

*The ANU College of Arts & Social Sciences*

Prof. Richard Rigby

*The ANU College of Asia & the Pacific*

Dr Maria Rost Rublee

*The ANU College of Arts & Social Sciences*

Dr Amelia Simpson

*The ANU College of Law*

Prof. William Tow

*The ANU College of Asia & the Pacific*

Ms Fionnuala Walsh

*Trinity College Dublin, Department of History*

Dr Xhengdao Ye

*The ANU College of Arts & Social Sciences*

Dr Ioannis Zoigas

*The ANU College of Arts & Social Sciences*
Cover Art

Estelita Rae

Bachelor of Visual Arts/Bachelor of Music

Leaf Pattern no. 6, 2013

watercolour on paper

14 x 19 cm

Photographer: Brenton McGeachie

My paintings centre around the exploration of growth patterns found in nature with specific references to flowers and leaves. I allude to patterns and symmetry in representation of a natural beauty found in these elements, as well as finding rhythms to express an understated harmony in the balance of nature.

The patterns in leaves and petals are delicate and fine, which unlike the sweeping mass of plant matter around us, allow us to take a closer to look at the finer intricacies of nature.

The palette I use is a derivative of collected source material and the general wash of colour I register around myself. It is also taken from details of leaves and flowers I have found, which I distorted to fulfil my own interpretation of the subject.

With this I hope to portray elements in nature such as flowers, leaves and water. I hope to capture the intricacies of these specimens and evoke a sense of remembering nature, whether it be the smell of flowers or the texture of the petals while holding the delicate membrane.
This year I have been looking at the concept of human-implied presence. I sought to create a work that implies human form in a simplistic and subtle way so as to emanate a presence. This comes from previous studies on how an art form can replicate human presence and the way in which humans leave a presence in the world through marks or imprints.

I additionally aimed to create a sculpture piece that crosses boundaries between mediums and dimensions, particularly focusing on how I can apply drawing to a sculptural 3D medium. This work consists of a series of ink drawings on muslin fabric that are supported midair by wooden frames. I chose muslin cloth because of its dual ability as both a sculptural and print medium. Throughout the past semester I have explored the possibilities of using these characteristics of muslin fabric with my interest in the human form/figure.
The Breaking of the ‘Great Australian Silence’: How and Why the Writing of Indigenous Australian History has Changed over the Last 40 Years

Caroline Beasley

Abstract

Australian history has not always recognised the traditional Indigenous occupants of the country. For various reasons, including prevailing ideas about race, Indigenous peoples were excluded from Australian history. Anthropologist WEH Stanner’s 1968 Boyer Lectures called historians to ensure this pervasive forgetfulness of the Indigenous population ceased, a process already under way when the lectures were delivered. The lectures had their intended effect of further encouraging the historical revision that subsequently altered the foundations of Australian history by creating a new field: Indigenous Australian history. These changes in historiography occurred gradually, and included many historians. The prevailing social and cultural forces of the time were also instrumental in aiding this revision, and included the 1967 referendum and subsequent land rights movement, societal tensions over the bicentenary celebrations in 1988, the High Court decisions of Mabo and Wik, as well as the political leadership of the time with which this field became intertwined. Ultimately Indigenous Australians moved from being a ‘melancholy footnote’ in Australian history, to occupying a central place in the historiography.

Introduction

It is not just a matter of attaching Aboriginal history to the back left hand corner of the old homestead or of even glassing in the back verandah. The changes will ultimately have to be far more radical — a new floor perhaps, even new foundations.¹

Henry Reynolds

The last 40 years have been a time of great change for the discipline of Indigenous Australian history. As Reynolds’s remark suggests, this change required a dramatic shift in how Australian history was conceptualised, such that ‘new foundations’ were mandatory for the traditional Australian historical discourse from which Indigenous Australians were excluded. Following anthropologist WEH Stanner’s 1968 Boyer Lectures, the scene was set for further historical revision of traditional ideas of Australia’s past, with Indigenous Australians in mind. As Stanner put it, ‘the native question is rising into great importance, the melancholy footnote is turning into a whole chapter of Australian history, and the codicil is becoming a major theme in the Australian story’.2

Many historians were part of this revisionism, which has evolved significantly since the 1970s, occurring incrementally and not without criticism and re-revision. These changes in the Australian historical discourse leading to greater inclusion of Indigenous Australians can be attributed to several factors including the political disposition of the time, landmark constitutional and legal changes such as the 1967 referendum, the land rights movement, public tensions over the bicentenary celebrations in 1988 and the Mabo and Wik decisions concerning Indigenous land rights, as well as the desire to create a national identity.3 The changes in Australian historical scholarship that have occurred over the last 40 years resulted in the stories of Indigenous people, and the Indigenous perspective, being increasingly included in the Australian historical narrative. This essay will focus on how ideas of Indigenous Australian history changed, before considering the factors contributing to why these changes occurred. First, however, it is necessary to explore what the phrase ‘Indigenous Australian history’ means.

What is Indigenous Australian History?

History itself is a primarily European discourse,4 the nature of which perpetuates the exclusion of Indigenous people from early Australian narratives.5 The then-dominant racial thought encouraged colonisers to understand Indigenous people

---

2 WEH Stanner, The Dreaming and Other Essays, Black Inc. Agenda, Melbourne, 2009, p. 176. ‘Melancholy footnote’ and ‘codicil’ were terms that had previously been used by historians and anthropologists to describe the status of Indigenous peoples in Australian history. See, for example, JA La Nauze, ‘The Study of Australian History 1929–1959’, Historical Studies Australia and New Zealand, vol. 9, no. 33, 1959, p. 11.

3 Mabo v Queensland (No. 2) (1992) 175 CLR 1; The Wik Peoples v Queensland (1996) 141 ALR 129.


as ‘preliterate’ and their society as ‘prehistorical’. This led to further affirmation of their exclusion from Australian history as they were unable to be the subjects of history. Any unfamiliar aspects of Indigenous life were assigned to ‘the fickle, wavering disposition of all savages’. Accordingly, the main subjects of traditional Australian history were the British, conceptualised as brave and resourceful pioneers who peacefully colonised the harsh land of Australia.

Given the difference between Indigenous and European cultures, a prominent issue for Indigenous history is the question of who is best placed to write it: if Indigenous history is primarily written by European historians, can it still be considered Indigenous history given its concern with Aboriginal experience? Some Indigenous writers, such as Jill Milroy, have labelled this dichotomy as the difference between writing Indigenous history, and history about Indigenous people. Although Indigenous people can offer unique insights into their history, some European historians have been active in the propagation of Indigenous history. These writers have often explicitly stated that their work is not an attempt to speak on behalf of Indigenous people, but rather a European interpretation, while others have made efforts to use Indigenous oral histories as sources to inform their work. As such, this essay defines Indigenous Australian history as those historical discourses in which Indigenous people are subjects, written either by Indigenous or non-Indigenous historians. It is worth noting that much of the scholarship in this field was created at a time when the field was known as Australian Aboriginal history. Indeed a prominent journal in this field is titled Aboriginal History and it aims to emphasise ‘not only … Aboriginal people as subjects and actors but Aboriginal perceptions and perspectives as well’. As such the terms ‘Indigenous’ and ‘Aboriginal’ are used interchangeably throughout the essay.

---

8 Stanner, The Dreaming, p. 93.
11 ibid., p. 2.
15 The author of this paper is a European Australian.
Changes in Indigenous Australian History

In the ‘most consequential lecture ever broadcast on the ABC’, Stanner highlighted how the Indigenous experience of colonisation and its effects had been increasingly omitted from the Australian historical discourse as developed in the twentieth century — the ‘great Australian silence’. This called historians to action:

what may have begun as a simple forgetting of other possible views turned into a habit and over time into something like a cult of forgetfulness practiced on a national scale.

This ‘cult’ was so pervasive that Australians struggled to keep the traditional occupants of the land in mind, even when it was important to do so, thus perpetuating the ‘great Australian silence’. At the time of his Boyer Lectures, however, Stanner noted that the Indigenous past was returning to public awareness. Indeed, since these lectures, historians have begun to create a ‘new Australian history’, with ‘many voices … fill[ing] out the space once claimed by [the] great Australian silence’. Since Stanner’s lecture, the process of change in Indigenous history has occurred incrementally, recognising different features of the Indigenous experience over time.

In the early 1970s historians began to write ‘social history’, which saw invasion as the main theme of colonisation, as opposed to the previously ‘peaceful’ arrival of the British that underscored traditional colonial history. The perceived incompatibility between Indigenous and European cultures was clear, especially surrounding land use, and there was an understanding that violence on the frontier had caused the British destruction of Indigenous societies. The work of this period saw Indigenous people as victims, which was a shift from the traditional historical discourse of ‘forgetfulness’ of the Indigenous population.

17 ibid.
18 Stanner, The Dreaming, p. 189.
19 ibid.
20 Attwood, Telling the Truth, p. 17.
25 Veracini, ‘A Prehistory’, pp. 441–42. CD Rowley was an important writer in this movement whose work predates Stanner’s lectures. By 1969 he had already advocated an Indigenous perspective in Aboriginals in Australian Society and, in 1972, published The Destruction of Aboriginal Society, which is an important example of the ‘social history’ of the 1970s.
In the late 1970s historians began to focus more on the scope of Indigenous resistance to the colonising force and the frontier was seen as ‘contested’.\textsuperscript{26} This also involved consideration of Indigenous perspectives, as pioneered by Reynolds in \textit{The Other Side of the Frontier} (1981).\textsuperscript{27} Indigenous authors also began to publish their own work at this time and this represented a significant change in the historiographical landscape, particularly given the personal nature of the stories and their basis in oral histories.\textsuperscript{28}

Influenced by Reynolds’s consideration of alternative perspectives, 1980s scholarship began to approach Indigenous oral sources as supplements to the written records.\textsuperscript{29} This era was also characterised by exploration of Indigenous agency.\textsuperscript{30} Ann McGrath’s research, which exemplified this through its focus on Aboriginal employment in the Northern Territory cattle industry, encouraged historical exploration of themes other than frontier violence.\textsuperscript{31}

More recently, scholarship on the nature of Australian frontiers has declined in favour of exploration of the possibility of genocide, as well as representations of Aboriginality.\textsuperscript{32} This has seen conservative historians mount an attack on the ‘black armband’ view of history. The term ‘black armband history’ was coined by historian Geoffrey Blainey in 1993 to describe history that focuses on the more sombre and negative aspects of Australia’s past, such as the treatment of Aboriginal people, rather than narratives of national achievement. According to Blainey, this approach was motivated by a desire to be politically correct, rather than by truth.\textsuperscript{33} Proponents of Blainey’s view advocated a return to traditional views of Australian history, a view labelled by revisionist historians as ‘white blindfold history’. The aim of this re-revision was to restore faith in the good character of the nation created by the British.\textsuperscript{34} This re-revision was inherently political.

Overall, since Stanner’s 1968 lectures, the great Australian silence has been filled with many voices, which satisfied Stanner’s call for an emphasis on Indigenous experience since colonisation.\textsuperscript{35} Indigenous people moved from ‘footnote’ to the forefront of Australian history, challenging its foundations and its structure.

\begin{itemize}
\item\textsuperscript{26} Veracini, “‘Contested Ground’”, p. 225; Veracini, ‘A Prehistory’, p. 443.
\item\textsuperscript{27} Attwood and Foster, ‘Introduction’, in Attwood and Foster (eds), \textit{Frontier Conflict}, pp. 6–7; Attwood, ‘Historiography on the Australian Frontier’, in Attwood and Foster (eds), \textit{Frontier Conflict}, p. 171.
\item\textsuperscript{28} Brock, ‘Skirmishes’, p. 211.
\item\textsuperscript{29} Attwood and Foster, ‘Introduction’, in Attwood and Foster (eds), \textit{Frontier Conflict}, pp. 8–9.
\item\textsuperscript{30} Veracini, ‘A Prehistory’, p. 448.
\item\textsuperscript{32} Attwood and Foster ‘Introduction’, in Attwood and Foster (eds), \textit{Frontier Conflict}, p. 10.
\item\textsuperscript{33} Brock, ‘Skirmishes’, p. 219.
\item\textsuperscript{34} ibid., p. 207.
\item\textsuperscript{35} Attwood, \textit{Telling the Truth}, p. 42.
\end{itemize}
Why has Indigenous Australian History Changed?

As Indigenous Australians became more prominent in historical discourses, these histories were increasingly linked with aspects of public life. The policies of contemporary governments, as well as political and legal endorsements of certain historical perspectives, contributed to the changes in Indigenous Australian history that have occurred since 1968. This was particularly true in the 1990s, when history’s increased politicisation ensured the ‘new foundations’ referred to by Reynolds were established within the Australian psyche.36

After the successful 1967 referendum which removed some constitutional discrimination against Indigenous peoples, Indigenous people were subject to Commonwealth legislation and able to be included in the census, a change that began the slow process of increasing legal rights and changing public perception. Reynolds vividly illustrated the pace of this change in a haunting depiction of a scene at a Palm Island jail in 1969, where two Indigenous schoolgirls were held as punishment for answering back to their kindergarten teacher.37 The referendum movement more broadly was significant as it resulted in a large section of the Australian community — over 90 per cent — mobilising in favour of the removal of discrimination against Indigenous peoples.38 The success of the referendum was also symbolically significant for Australia on an international level as, by the 1960s, numerous other settler societies had recognised equal rights for colonised peoples.39 Australia’s Indigenous peoples were able to seize the momentum embodied in the referendum movement to start a more organised and more widespread land rights movement. The Tent Embassy, an enduring symbol of this struggle, was established in Canberra in 1972 and aimed to secure Indigenous title to land, mining royalties and other compensation, as well as illustrating symbolic sovereignty over the land.40 The end of discriminatory immigration and assimilation policies, which detrimentally affected Indigenous people, occurred concurrently with the slow beginning of Australian historical revision in the early 1970s. This was encouraged by the prominence of the subsequent land rights movement.41

---

37 Reynolds, Why Weren’t We Told?, p. 9.
38 ibid.
39 Attwood, Telling the Truth, p. 20.
The notion of national identity is rooted in history, and the country’s or individual’s place within it.\(^{42}\) Traditionally, the Australian identity was defined by ‘Britishness’, but following World War II, this identity was challenged.\(^{43}\) Britain’s decision in the early 1960s to focus on its economic ties with Europe, when it applied — albeit unsuccessfully — to join the European Economic Community,\(^{44}\) increased colonial challenges to imperial rule.\(^{45}\) Changes in Australian immigration policies and increasingly interlinked cultures created by global economic practices gradually led to a desire to redefine ‘Australianness’\(^{46}\). White Australia now sometimes sought to differentiate itself by drawing on notions of Aboriginality and ancientness.\(^{47}\) The increased desire to appropriate Aboriginality for the purposes of national identity stemmed from scholarly revision of traditional notions of history, archaeological research that was pushing the original Aboriginal occupation of the continent back thousands of years from the date previously identified, and recognition in other settler societies of Indigenous rights.\(^{48}\) The increased prominence of Indigenous culture led to more widespread acceptance of, and even reverence for, Indigenous perspectives, which — while not reflecting negatively on all aspects of British colonisation — was demonstrated by changes in contemporary Australian historical discourses.\(^{49}\)

The commemoration of the bicentenary of Australian settlement in 1988 saw the beginnings of a re-revision of Indigenous Australian history and its intertwining with national politics, which climaxed in the years of John Howard’s Liberal government. In 1980 Prime Minister Malcolm Fraser created the Bicentennial Authority to determine the method of celebrating, advocating that the authority take a ‘warts and all approach’ to selecting ways to represent Australia’s past.\(^{50}\) The Uniting Church and other groups opposed a British-centric commemoration of this occasion, alleging that this would represent ‘the destruction of Aboriginal society’.\(^{51}\) The attempts of the authority to address the concerns of every group in Australian society in an inclusive way led to claims from conservative historians that they were trying to create ‘a guilt industry’, which was disrespectful to Australia’s British past.\(^{52}\) The tension between the

\(^{42}\) ibid., p. 30.
\(^{45}\) ibid., p. 20.
\(^{47}\) ibid., p. xxiii.
\(^{48}\) Attwood, Telling the Truth, p. 20.
\(^{50}\) Stuart Macintyre and Anna Clark, The History Wars, Melbourne University Press, Carlton, 2003, p. 97.
\(^{51}\) ibid., p. 98.
\(^{52}\) ibid., p. 109.
dramatic revision of Indigenous Australian history by Reynolds and others and the re-revision offered by conservative historians, was established and continued into the future following 1988.

The 1992 Mabo case and the Wik case in 1996 established native title land rights for Indigenous people, and overturned the key legal doctrine of terra nullius on which Australia’s colonisation was based.\(^{53}\) This doctrine provided that in 1788 Australia was a land belonging to no one, despite the resident Indigenous population, enabling the British to claim the land and creating a ‘lawless terrain’ of subsequent relations.\(^{54}\) The role of Mabo and, in particular, the judgment of Justice Deane and Justice Gaudron had ramifications for subsequent historical scholarship, since it clearly established the legal status of Indigenous people since 1788.\(^{55}\) Indeed, the dispossession of Indigenous land and heritage that had occurred after was described in the joint judgment, which relied heavily on historical work undertaken by historians, such as Reynolds, as ‘a national legacy of unutterable shame’.\(^{56}\) The Wik case explicitly established that native title could exist on land that was the subject of a pastoral lease.\(^{57}\) The two cases led to an increase in writing of local histories, and emphasis on Indigenous agency as part of the investigation of communities’ ties to land to prove native title.\(^{58}\)

The legal changes that resulted from Mabo and Wik also encouraged increased politicisation of history due to the immense challenges posed to traditional notions of history, and thus to Australian nationhood and identity.\(^{59}\)

The relationship between historical scholarship and political shifts became more explicit in the 1990s following Mabo.\(^{60}\) As Labor Prime Minister in the 1990s, Paul Keating illustrated his support of the revisionist view of Indigenous Australian history, and the finding in Mabo, in a speech at Redfern Park, which was written by his speechwriter, historian Don Watson. The speech recognised, and called upon the nation to recognise, the other side of much of Australia’s history since colonisation:

> it was we who did the dispossessing. We took the traditional lands and smashed the traditional way of life. We brought the diseases. The alcohol. We committed the murders. We took the children from their mothers. We practised discrimination and exclusion … \(^{61}\)

---

53 Veracini, ‘‘Contested Ground’’, p. 226.
54 ibid.
56 Mabo v Queensland [no 2] (1992) 175 CLR 1, 104.
58 Veracini, ‘‘Contested Ground’’, p. 226.
60 Veracini, ‘‘Contested Ground’’, p. 224.
This explicit endorsement of the revisionist view, which amounted to an apology for past injustices committed by Europeans against Indigenous people, met its ideological and political opponent in the Howard government, elected in 1996. Howard publicly used Australian history as a ‘political weapon’, which due to his public status, had a strong impact on perceptions of Indigenous Australian history. Howard condemned the ‘politically correct’ revisionist historians — or the ‘black armband brigade’ — because he believed they postulated a version of history which made Australians ashamed of their past. Strongly advocating the return of patriotism, Howard stated that the Australian nation should not have to apologise for the actions of their ancestors, and laid blame on the revisionist historians for labelling the work of British settlers as insignificant. This reignited some of the historical debates that had occurred around the celebration of the bicentenary of Australian settlement a decade earlier. Howard’s advocacy of the single, patriotic, traditional narrative of Australia, which had rigorously defined perceptions of Indigenous sovereignty and status, led to his indirect support of the conservative historians through public print media. Conservative magazine Quadrant launched several campaigns in accordance with Howard’s ‘white blindfold’ view. The historian Keith Windschuttle was heavily involved in this and was often labelled as ‘Howard’s historian’. His work, which sought to dispel the ‘myths’ advocated by revisionist historians in the field of Indigenous Australian history, and pushed for a return to the ‘traditional’ version of Australian history, precipitated a historiographical battle known as the ‘history wars’. Despite some serious public consideration of Windschuttle’s points, the ‘black armband’ view remained ingrained in the national consciousness, and was not easily forgotten. As such the narratives of denialism which became prominent during this period and represented a shift away from the histories published since 1968, were considered, but not in such a way as to create a new orthodoxy.

Following Mabo, Indigenous dispossession became a recognised theme in Australian history and, as such, was a key part of the Australian Human Rights and Equal Opportunity Commission’s (HREOC) 1997 Bringing Them Home report on the Stolen Generations. The report relied on testimonies from individuals affected by child-removal policies, illustrating the importance of memory and Indigenous oral histories in creating history, despite condemnation of these
'sources' by the Howard government. The Bringing Them Home report is another example of the politicisation of Indigenous Australian history, and the effect this has had on the public. A key recommendation of the report was an official government apology to the members of the Stolen Generations, which Howard refused to make. Yet, due to the prominence of Indigenous history in public life, this decision was widely condemned. The public condemnation of Howard’s decision not to apologise to Australia’s Indigenous people illustrates clearly the changes that occurred in the field of Indigenous Australian history since 1968 — the new foundations for the history of Australia were now firmly in place.

Conclusion

This essay has illustrated the changes in Indigenous Australian history and why they have occurred over the last 40 or so years, from when Indigenous people were largely forgotten in the Australian historical discourse, to their essential presence in notions of Australian identity. Concomitantly, legal and political changes, including the 1967 referendum, subsequent land rights movement and High Court decisions, ensured that the historical changes were reflected in policy and in the settler Australian psyche, where the presence of Indigenous Australians as subjects and writers of history is now a permanent feature. This permanency is encouraged by the myriad alternative sources from which to gain historical knowledge, including novels, radio, art and songs. The increased democratisation of Indigenous Australian history is especially important; due to the increasing numbers of Indigenous and non-Indigenous contributors, consumers of history have a way to understand historical events from many perspectives, something that is likely to characterise the field of Indigenous Australian history into the future and facilitate its further expansion.

Bibliography


72 Attwood, Telling the Truth, p. 51.
73 Manne, Introductory’ in Manne (ed.) Whitewash, p. 4.


Davidson, Robyn, ‘No Fixed Address: Nomads and the Fate of the Planet’, Quarterly Essay, no. 24, 2006.


McKenna, Mark, Looking for Blackfellas’ Point: An Australian History of Place, University of New South Wales Press, Sydney, 2002.


Heritability Estimates of Complex Intelligence and Associated Genetics

Nyomi Bourbous

Abstract

It has always aroused curiosity in the field of human science studies as to how much of ourselves can be genetically attributable to our parents. For relatively simple traits, like earwax consistency and albinism, this poses no problem as they can be followed genetically using simple Mendelian methods. When it comes to more complex constructs, however, such as cognitive capabilities, it is difficult to determine, to say the least. Intelligence is one facet of human cognition that inspires awe and curiosity among scientists. Whether a person is born with a particular intelligence measure (nature), or if intelligence is developed through a nurturing and stimulating environment (nurture) is an ongoing debate. Evidence exists for both sides of the argument, but the most compelling comes from the use of genetically identical individuals whose variance in intelligence can perhaps answer questions regarding the impact of one’s surroundings on this complex cognitive capacity.

Keywords


Introduction

It has been shown that certain physical characteristics, such as finger lengths, correlate strongly with behaviours and personality (Paul et al., 2006). Thus it would make sense to assume that intelligence is also inherited. This assumption, however, underestimates the complexity of the nature of intelligence. When researching the heritability of intelligence, several factors come into play. Firstly, how can intelligence be quantified? What tests, definition, skills and traits make someone more intelligent than another? Is there a gene (or a set of genes) that localises intelligence to specific regions in the genome? Finally, it is imperative
to discuss the nature versus nurture debate; are people born smart or do they become intelligent through upbringing and exposure to the right environment? The debate surrounding whether or not intelligence is inherited, much like physical characteristics, is nothing new. This essay shall attempt to answer all these questions as well as examine the validity of intelligence as a component of genetics and the environment.

What is Intelligence?

Over the centuries, the subjective nature of what constitutes intelligence has given rise to various theories in the philosophical and scientific communities. The two main models, which have been at the forefront and therefore heavily debated, are the theory of multiple intelligences (MI) and the theory of general cognitive ability, denoted by the factor ‘g’ (Deary et al., 2009).

In 1986 Gardner expanded upon the Primary Mental Abilities Theory (formally established in 1938 by Thurstone) with the design of his own theory; that of Multiple Intelligence (Morgan, 1996). This theory stipulates that there are eight categories (not necessarily numerically based) for which one may display performance exceeding that of the general population, or in other words, the average. These include linguistic, logical-mathematical, musical, spatial, bodily kinaesthetic, intrapersonal, interpersonal and naturalistic intelligence (Traub, 1998). If an individual exceeds the ability of the general population in one of these fields of expertise, then that individual is deemed intelligent in that specific field. Gardner asserts that intelligence should not be determined by a single intellectual aptitude, but rather on a scope of independent variants (Traub, 1998).

In light of this account, Deary et al. (2009) claim that the MI theory holds no leverage in comparison with Charles Spearman’s 1916 theory of general cognitive ability, which accounts for 40–50 per cent of intellectual variance. This model works by calculating a numerical value of mental aptitude, the *general cognitive factor* or *g*, using the technique of factor analysis across a variety of tests. Spearman concluded that individuals who performed well on one cognitive test generally did so across the remainder. Conversely, if a poor result was obtained this was also reflected in the others (Kuncel et al., 2004). Ultimately the *g* factor represents a general level of intelligence that is applicable to everyone. It is when one displays a greater intellectual capacity than this average that intelligence can be noted. Although there is some debate regarding the specific test variances giving rise to *g*, these are irrelevant as it is principally *g* that carries the predictive validity as well as variance of the intelligence trait (Deary et al., 2009).
A variety of tests exist today that attempt to measure and quantify intelligence. Derived from the earlier Binet–Simon Scale, the Stanford–Binet Intelligence Quotient (IQ) Test founded in 1916 is considered to be one of the most valid tests. Debate continues with regard to its validity based upon cultural biases, gender differences and heritability. Most modern forms of the intelligence test have been specifically modified to remove as much of these as possible and instead compare individuals within their own age group (Bartholomew, 2004).

As is clear, there are a variety of approaches to understanding and measuring intelligence. Yet the underlying mechanisms that enable the existence of the enormous variety of intellectual traits are still unknown.

The Heritability of Complex Traits

In order to assess whether a genetic link exists between genes and intelligence, it is important to determine what genetic heritability is and whether it is a causal factor for assuming characteristics from one’s parents. Heritability is recognised as the degree of phenotypic variation for a given trait resulting from genetic variation (Mackintosh, 1998). When it comes to simple Mendelian traits, such as earwax consistency and albinism, it is easy to distinguish whether genetics play a significant role in the phenotype of an individual as it simply requires the presence of a given gene. Yet cognitive abilities, like intelligence, are complex traits in which many genes accounting for small effects are hypothesised to contribute to their variation (Luciano et al., 2006). For this reason they are difficult to study. For example, to find a gene responsible for haemophilia, it is only necessary to look at a family history and extrapolate the genotypes from the phenotypes displayed throughout the pedigree. But with complex traits, multiple genes interact and are differentially expressed in individuals. Thus it is difficult to pinpoint where in the genome such genes exist.

The Heritability of Intelligence

For centuries people have contemplated whether a link exists between genes and intelligence. As technology and communication increase we are seeing more cases of siblings, such as the members of the British Imafidon family (Campbell, 2009), arise whereby multiple individuals within families exhibit superior intellectual capacities, which would not occur through random variation. Such examples lead to the question of whether intelligence has a genetic basis.

The development of the ability to map the human genome has enabled scientists to map specific genes that may relate to intelligence. The two main strategies
in finding such genes include linkage analysis and candidate-gene association (Posthuma & Geus, 2006). Linkage analysis is a method by which multiple markers affixed to known locations are correlated with the presence of certain phenotypes. If genes are found in close proximity to such markers then it makes it easier to locate them. This is because if the inheritance of a marker and trait can sufficiently be proven to occur together, ‘then one or more genes in the vicinity of the marker are possibly involved in the trait similarity because genetic similarity on chromosomes is passed on in chunks’ (Posthuma & Geus, 2006). Conversely, candidate-gene association has been described as the process in which genes previously associated with a characteristic or a trait are tested to discern whether a relationship exists between them (Patnala et al., 2013). This is done by comparing the gene frequency of individuals who are known to express the phenotype against control subjects; i.e., individuals who do not express the trait. Association is inferred if the gene is observed at a significantly greater frequency in individuals expressing the trait (over control subjects).

It is important to note, however, that both these techniques are relatively new. While only a limited amount of research has been performed on their accuracy, studies were conducted, specifically in 2005 and 2006, that used linkage analysis on intelligence (Buyske et al., 2006; Dick et al., 2006; Luciano et al., 2006; Posthuma et al., 2005; Wainwright et al., 2006). The information gathered from these studies suggests that there are multiple genes, spanning 15 chromosomes (out of the 24, including x and y as separate units), that may contribute to intelligence (see Figure 1). More specifically it is unknown if and how these genes are related to intelligence even though many of them are found to be weakly correlated with the phenotypes that make up intelligence. Two regions on chromosomes 2 and 6 have, however, been positively identified as contributing to intelligence (Luciano et al. 2006; Posthuma et al., 2005).

Research using model-free multipoint linkage analysis has determined that region q24.1-31.1 of chromosome 2 and region p25.3-22.3 of chromosome 6 appear to display high correlations with a quantitative-trait locus for performance (perceptual organisation) intelligence quotient (IQ) and a suggestive linkage for both full-scale (cognitive capacity) and verbal (verbal comprehension) IQ respectively (Posthuma et al., 2005). These regions were identified because they presented a high correspondence between intellect and certain markers. The region of chromosome 2 highlighted an association with performance IQ (PIQ) and the marker D2S2330, while chromosome 6 indicated an association between full scale IQ (FSIQ) and verbal IQ (VIQ) via the marker D6S942.

Even more interesting is that each of these regions has been connected with various mental disorders (Posthuma et al., 2005). The q24.1-31.1 region of chromosome 2 largely overlaps with the 2q21-33 region, which has been strongly linked to autism by at least four independent genetic studies (Philippe et al.,
1999; Buxbaum et al., 2001; Palferman et al., 2001; Shao et al., 2002). Likewise, the p25.3-22.3 region of chromosome 6 also displays an overlapping region that is related to reading disabilities and dyslexia. Most intriguingly, several genes exist within these linkage regions that have been associated with schizophrenia, fragile-X syndrome and Bardet-Biedl syndrome — disorders that are typically accompanied by cognitive impairments.

Figure 1: An ideogram of the human genome (based on the five linkage studies of intelligence conducted to date) highlighting chromosomal regions (denoted by orange) containing genes already associated with intelligence and regions (denoted by red) believed to contain genes associated with intelligence.


A 2006 study (Luciano et al.) reported similar results to those described above (Posthuma et al., 2005). The study found that there are regions in chromosomes 2 and 6 that correlate with PIQ and FSIQ/VIQ respectively. It also demonstrated that the linkage region of chromosome 2 not only correlated for PIQ, but also for word recognition. This confirmed that there could be individual genes that influence different aspects of cognitive ability (Plomin & Kovas, 2005).

Genetic analysis provides an understanding of the correlations between genes and phenotypes through testing IQ scores and comparing them to genes. This does not, however, test for the interactions between nature and nurture. Indeed,
epigenetics further complicates the connection between intelligence and genetic heritability. Twin studies have been used in attempts to separate the effect of the environment from that of the genes.

The Relevance of Twin Studies and Adoption Studies

Twin studies and adoption studies have been undertaken since the late nineteenth century, when they were first utilised by Sir Francis Galton (Yong-Kyu, 2009). Galton surmised that the more genetically alike two individuals are, the more similar their physical and mental characteristics.

Monozygotic twins share 100 per cent of their genes while fraternal twins are exactly like individually born siblings and share approximately 50 per cent of their genes. It logically follows then that ‘if identical twins are more similar than fraternal twins on a particular outcome, then genetic influences are assumed’ (Wilhelm & Engle, 2005). Likewise, if adopted children are more similar to their adoptive parents and siblings, then environmental influences are implicated. Twin and adoptive studies are thus frequently used in genetic research as means of ascertaining environmental and genetic influences on different traits.

Such studies have advanced genetic research. Unlike rare, clinical conditions, which can be diagnosed by observing a regular pattern of inheritance throughout a family’s history, complex traits cannot be measured in this way (Johnson et al., 2009). As discussed, there are multiple genes that interact to magnify or reduce the effect of other genes. Twin studies allow researchers to effectively separate the gene’s effects from the environment. Monozygotic twins reared apart are genetically identical individuals living in different environments. By examining the differences that exist between them in particular characteristics (e.g., IQ scores), it is possible to ascertain the effect of the environment. The downside to these studies is that they require large sample sizes from monozygotic twin populations (which are rare) in order to accumulate sufficient statistical power (Johnson et al., 2009).

The Nature versus Nurture Debate

At the forefront of human genetic research exists the debate as to whether intelligence is a preconceived trait or one acquired over an individual’s lifetime. Although when first established in the 1970s the debate consisted of two clear unambiguous positions, that of nature versus nurture, more recently it
has moved on to the theme of nature via nurture (Anderson, 1999). Nature via nurture means that both components contribute differently to intelligence and that, symbiotically, one cannot work without the other.

Today there is no denying that the heritability of intelligence is one of the most consistent findings in the scientific literature (Wilhelm & Engle, 2005), which has been determined mainly by twin and adoption studies (Deary et al., 2009). These methods have found significant correlations between age and pronounced intellectual heritability as well as socio-economic status and intelligence. Unlike Mendelian genetics, intelligence does not exhibit a uniform hereditary status throughout one’s lifetime. Instead, with the assistance of twin studies, it has been found that generally, heritability of intelligence increases from around 20 per cent in infancy to 60 per cent in adulthood onward to 80 per cent in old age (Wilhelm & Engle, 2005). This information reveals that nurture is significant to the growth and developmental expression of our genes. Similarly, analysis of the heritability of intelligence as impacted by socio-economic upbringing has found that individuals raised in low socio-economic environments display low heritability of intelligence while individuals raised in high socio-economic environments display increased levels of heritability (Posthuma & Geus, 2006). This also indicates that nurture, via our environment, is influential in the expression of intelligence.

Ultimately with the assistance of twin studies and new genetic technology the nurture–nature debate is leaning in the direction that both components are significant to the development of intelligence. The naturalistic components are the genetic contributors and it is acknowledged that genes form the basis upon which intelligence can be established and that everybody possesses a basic intellectual capacity (denoted by the value g). Also, the component of nurture is required to further enhance an individual’s intelligence, which involves the development of neurological pathways and thought processes. Yet it is the maturation of these cognitive processes that sees them switch from being environmentally activated to having a genetic basis.

It can be concluded that, although nature is a necessary component to providing a foundation upon which intelligence can grow, nurture is just as crucial to continue building upon this foundation, for it appears that an environment void of intellectual stimuli would inhibit this growth.
Possible Future Directions for the Study of Intelligence

The field of genetic inheritance of intelligence is recently established, mainly because the mapping of the human genome concluded only fairly recently, especially when compared to the field of intelligence and its heritability as a whole. As this essay has demonstrated, new research has deduced that intelligence does indeed have a genetic basis, with specific areas in chromosomes 2 and 6 possibly being influential. Although this is the case, researchers are still unable to infer how these areas contribute. Future investigation of this relationship will assist the medical world to understand the development of intelligence and the biological meaning of its apparently high heritability (Deary et al., 2009). Additionally, it would open new windows into the investigation of mental disorders and their diagnosis, as well as possible treatments and cures.

With regard to the influence of nurture on the expression of intelligence, it has been established that everybody possesses a basic level of intelligence. Heritability is, however, increasingly expressed with age and as a result of an advantageous socio-economic upbringing. This suggests that the greater the development of the individual’s brain, whether as a result of growth or external stimuli, the more intelligence functions will begin to resemble those of the parents.

Another promising route for future research into intelligence is the study of knockout genes. This method attempts to ascertain whether an individual exhibits a faulty gene (or lack of it) and whether its absence impairs their intellectual aptitude. Genetic knockout experiments on humans would give rise to many ethical objections, which may be overcome by studying individuals who already exhibit these mutations and by conducting animal model experiments.

In conclusion, it is clear that genes and an individual’s environment influence intelligence. To place an exact number or percentage on this influence would, however, be nearly impossible as there are multiple factors that determine the proportional influence of each.

Bibliography


Tradition and the Art of Modern India

Kieran Browne

Notions of tradition are complicated when discussed in relation to non-Western modernisms, such as that of India. Traditions resisting change have either been reinvented and modernised without consent, or abandoned in favour of those that are applicable to a modern world. KG Subramanyan is one artist who has moved to assimilate the traditional and the modern in a twentieth century Indian context. His work appropriates elements of Indian visual tradition, as well as popular and classical elements and reinterprets them through modernist forms and techniques. His works Girl with pigeons (1959–1960), Woman with lamp 1 (1951) and Industry (1955) inherit global visual culture, containing elements of Hindu iconography, Euro-American modernism and classical fresco painting. Subramanyan characterises a key section of the broader Indian modernist movement, which seeks to redefine tradition as a living, changing language. This essay argues that the interaction of tradition and modernity in India is indicative of a broader reality: that tradition is a fluid, living form.

Tradition is in flux, finding its ideas and practices constantly reinvented to suit present purposes. In the development of non-Western modernisms, such as that of India, visual traditions have collided with external elements, such as those from Europe. Unlike the Western avant-garde, which has typically met tradition with hostility, Indian modernist artists have moved to assimilate the traditional and the modern, reinventing both in a twentieth century Indian context.1 The role of tradition in modern culture is a key concern for Indian modernists and has been explored by influential Indian artists and groups including Rabindranath Tagore, Nandalal Bose and the Bombay Progressives. This essay examines the work of Subramanyan, one of India’s most influential artists and educators.2 Subramanyan’s examination of tradition, both as an artist and art theorist, has influenced many distinguished artists and made a clear imprint on Indian modern and contemporary art.3 This consideration of Subramanyan’s work explores his background and influences and the expression of tradition in his work through religious iconography and global influences. The essay concludes by examining contradictions in the notion of tradition, and the need for reinvention in order to sustain culture.

‘Tradition’ is difficult to define in relation to non-Western modernisms. The dichotomy of the traditional and modern, as it is often understood in the West, is not transferable to Indian culture. Tradition can be defined as the craft and language of culture: ‘craft’ refers to the methods, materials and tools used in any creative endeavour while ‘language’ refers to the images and ideas that inform the creator. By this definition, modernism is itself a tradition with its own craft and language. For this reason, in order to explore the relationship between the modern and the traditional, it is useful to refer to pre-modern elements of Indian artistic tradition, such as Hindu and Mughal history, as ‘indigenous tradition’.

Subramanyan was born in Kerala, south-west India, in 1924. As was typical of the region, he received a good education and had ready access to a French-administered library where he discovered the writings of Indian intellectuals and, amongst many French magazines, The Modern Review.4 As a young man he was involved in a number of political movements and led student protests during his time at university in Madras, where he spent six months in jail for his activism. On Subramanyan’s release in 1943, Bose, then head of Visva-Bharati University, Santiniketan, invited him to study at the university. There, Subramanyan developed an ideological framework for his artistic practice that was profoundly influenced by Bose as well as Tagore, the university’s founder.5 The doctrine that pervaded teaching at Visva-Bharati was that tradition informs community, which in turn nurtures individuals and culminates in tradition being advanced by an individual’s artistic practice.6 Subramanyan understood artistic and visual traditions of India and the wider world; he believed artists must abbreviate and reinvent the elements of the visual language they consume.7 Influenced by popular, modern, classical and indigenous traditions, Subramanyan’s artistic practice was concerned with assimilating these fluid traditions and reinventing them for present purposes.

The fluidity of tradition can be seen in the integration of modernism and indigenous tradition that occurred in India in the twentieth century. For many Indian modernist artists, Hindu iconography was an important part of the Indian artistic tradition, which their work revived and re-imagined during this period.8 The use of Hindu iconography for artists at the time was as much a visual form of secularism as an aesthetic decision.9 At that time there was widespread use of religious imagery by Hindu nationalists as a symbol of Hindu cultural and

---

5 Kapur, Subramanyan, 18.
6 ibid.
7 Kapur, Subramanyan, 3.
8 Zitewitz, ‘Secular Icon,’ 12.
9 ibid.
political domination. This political trend, which excluded non-Hindus, was a cause for concern amongst intellectuals and prompted many modernist artists to reclaim the icons through a series of secular appropriations.10

![Image](image_url)

**Figure 1: Girl with pigeons, KG Subramanyan, 1959–1960**

Source: Geeta Kapur, *KG Subramanyan* (New Delhi: Lalit Kala Akademi, 1987), 27.11

They argued that Hindu iconography belonged to Indian visual culture generally and could be reinterpreted as any other piece of culture. In line with this, Subramanyan believed that all faiths and institutions should be subject to criticism as an essential part of culture. He viewed visual culture as an archive of images and ideas available to artists for citation.12 Figures 1 and 2 exemplify the way in which Subramanyan references two Hindu icons within identifiably modern painting styles.

10 ibid.
11 The images in this article are reproduced from Kapur’s monograph on Subramanyan, in which the reproductions are black and white.
Subramanyan’s oil painting *Girl with pigeons* (Figure 1) is a prime example of the assimilation of modernist and indigenous traditional elements. Painted between 1959 and 1960, the work displays the modernist simplification of forms and flattening of the image plane, its visual style appearing to draw from European modernist painters, particularly Henri Matisse. These modern elements are incorporated with elements of indigenous visual language. The painting cites the Hindu deity Rati, goddess of love, passion and sexual pleasure. While Subramanyan maintains distance from the religious potency of the icon, he simultaneously comments on and represents it. The painting imagines a modern iteration of Rati in the form of a young woman in the street. She is not dressed as a deity nor does she stand as one, however, the reference is made clear by the presence of the pigeons, the symbol of Rati.

The purpose of Subramanyan’s reference to the Hindu icon is twofold: firstly, it allows him to quickly communicate a representation of the girl as an object of sexual desire, drawing on the collective Indian understanding of its dominant Hindu culture; secondly and more importantly, Subramanyan is reclaiming Rati. He transplants the icon into a modern, secular context and, in doing so, reinvents both modern and indigenous tradition for a present purpose. Subramanyan saw his work as an interpenetration of religious content; an act which enlivened those icons rather than damaging them. Subramanyan shows that the elements of Indian indigenous tradition are in flux, their craft and language being reinvented rather than cast away.

The fluidity of tradition is highlighted in the modern period through the assimilation of modern, classical and popular visual cultures. This development was nowhere more apparent than in India, where many believed that the modernisation of indigenous traditional art was vital to its preservation. Tagore argued that Asia was defined by its tradition and should modernise ‘on its own terms’. As a universalist and a humanist, who saw himself as a custodian of world culture, Tagore had no qualms in combining modernist, popular, classical and indigenous traditions in the modern context. He declared that ‘all traditional structures of art must have a sufficient degree of elasticity.’ His artistic philosophy was outlined in the essay ‘Art and Tradition’ in which he urged Indian artists to deny their obligation to create anything that aligns with

---

13 Zitzewitz, ‘Secular Icon’, 16.
15 Zitzewitz, ‘Secular Icon’, 17, 22.
the ‘old world’ idea of Indian art. Tagore’s beliefs reflected those held by many Indian modernists: that the aim of modernism was not to suppress tradition, but to widen its bounds and create a single, global visual language.

Profoundly influenced by Tagore’s artistic practice, Subramanyan engaged with Tagore’s eclecticism. Subramanyan believed multivalence to be endemic across Indian indigenous artistic tradition and a characteristic that was definitively Indian; he aimed for multivalence in his work to ensure that his meaning was left open to interpretation. Subramanyan’s years at Visva-Bharati solidified for him the importance of drawing upon and reinventing tradition. Being well-travelled, well-read and having a good understanding of Indian as well as non-Indian cultures, Subramanyan was skilled at marrying the modernist, popular and classical traditions with the indigenous. His work illustrates how elements of modern and classical traditions can inform and reform an artist’s practice.

Figure 2: Woman with lamp 1, KG Subramanyan, 1951

Source: Kapur, Subramanyan, 24.

22 Kapur, Subramanyan, 3.
Woman with lamp 1 (Figure 2) is one of the many Western modernist experiments carried out by Subramanyan in the 1950s. The influence of European modernist artists is evident in the work, particularly Matisse and Pablo Picasso. The influence of Matisse is expressed in the painting’s form and use of colour. Matisse’s work strove for ‘serenity’ through simple contours and forms, and Subramanyan reflects this ideal in his use of solid blocks of flat colour, which dispense with volume and indicate form through line and shape instead. In Woman with lamp 1 colour blocks are stacked upon one another from light to dark to create the illusion of shadows and highlights. The simplification of form and flattening of the image plane is common to many early European modernist artworks and translates well into Subramanyan’s art practice. References to Picasso can be seen in the multiple perspectives and dissection of the figure and her surroundings. In the image, a number of perspectives have been condensed into the image plane. The figure’s face is represented in portrait, but the rounded form in the lower colour block implies a full-face view. The influences from European modernism on Subramanyan’s work, and the ease with which these traditions are combined, can be seen as evidence of the fluidity of tradition. Subramanyan, in assimilating diverse traditions into his own practice, has reinvented modernism within a present, Indian context. In doing so he has strengthened and enlivened the Indian artistic tradition.

Industry (Figure 3) displays the integration of both classical and modern traditions. The subject matter is modern and placed in an Indian setting, but it is arranged in the style of a classical Greek or Roman fresco. The painting displays the characteristic compression of three-dimensional space that is typical of classical frescoes. The figures are stacked on top of one another to indicate depth and there is no visual perspective; the faces of each brick column are flattened, apparently perpendicular to the viewer’s gaze. Many of Subramanyan’s teachers at Visva-Bharati were interested in fresco painting and this is most likely the influence for the classical layout in this image. Modernist influences are also apparent in this painting, the theme of which is modernisation. These influences can be seen in Subramanyan’s complex use of the gaze in this work. Initially the viewer’s gaze is directed to a flattened perspective, which gives the impression of the image being viewed directly. More intriguing, however, is the use of the gaze amongst the figures, the majority of whom stand, absorbed in their task of observing the viewer. The figures are depicted with a wide, unsettling stare and an expressionless face. The diverse modern and classical elements in this painting have been smoothly translated into an Indian setting and Subramanyan has utilised the fluidity of tradition, reforming his artistic practice and reinventing these traditional elements for his present purposes.

26 Kapur, Subramanyan, 23.
Figure 3: Industry, KG Subramanyan, 1955

Source: Kapur, Subramanyan, 23.
Tradition is a concept in constant contradiction. It is the basis on which a society and culture preserves its selfhood. Yet it must be constantly reinvented in order to be kept ‘alive’. JC Heesterman describes this contradiction aptly, writing that tradition aspires to reach some fundamental and yet unreachable conclusion.\(^{27}\) It is this pursuit that keeps tradition in flux, yet firmly connected with the present. The conflict of tradition arises from a long-standing debate; does the reinvention of tradition preserve or undermine it? Tradition is often misunderstood as a cultural preserve. Nevertheless, while a society is concerned with maintaining its selfhood, it is also concerned with remaining a living, growing and evolving organism. Following this argument, elements of tradition that are not in present use are in decay. Tradition must be in continual use in order to be sustained.

Subramanyan’s artistic and intellectual pursuits in the modern period have been concerned with the role of tradition in modern Indian culture. His experiments with interpenetration of Hindu iconography and the assimilation of popular, modern, classical and indigenous traditions have been vital to the revival of Indian visual culture. In continuing the work of Tagore and Bose, Subramanyan has enlivened Indian indigenous tradition and reinvented it for present purposes. His examination of tradition has influenced much of India’s modern and contemporary art. As India undergoes rapid industrial modernisation, the question of whether culture ought to embrace modernism, or hold on to the past, inevitably arises. Traditions resisting change have either been reinvented and modernised without consent, or abandoned in favour of traditions that are applicable to an increasingly modern world. Modernism is not a suppression or supersession of tradition. Rather, it is a tradition by itself in a global context. The assimilation of modern and indigenous traditional elements allows tradition to fulfil its present purpose creating a global, modern visual culture. Tradition in modern India is alive, sustained by the work of artists such as Subramanyan. It is, if not by definition then by necessity, a fluid entity.

**Bibliography**


---

\(^{27}\) Heesterman, ‘India’, 97–99.


The Promise and Failure of King Amanullah’s Modernisation Program in Afghanistan

Andrew Chua

Abstract

Nearly a century before the United States embarked on the perilous task of transforming Afghanistan into a modern nation state, King Amanullah dreamed of his country joining the ranks of modern states and embarked on a wide-ranging modernisation program. For ten years no area in society was left untouched by Amanullah’s reforms; which, had they succeeded, would have transformed Afghanistan into a modern and westernised nation state. By 1929, however, the self-described revolutionary had been forced into exile, never to return to his beloved country, and his reforms were abrogated. This paper examines the policies and objectives of Amanullah’s modernisation program. It argues that Amanullah mismanaged the reform process to the point of alienating almost every group in society. This led to a tribal revolt that brought an end to both Amanullah’s reign and his modernisation program.

Introduction

More than ten years after invading Afghanistan to defeat the Taliban and establish a modern nation state, the United States finds itself unable to achieve those objectives and it has had only limited success in the Afghan nation-building effort.1 The current situation in Afghanistan mirrors another episode in the nation’s past — the modernisation program of King Amanullah who, over a century earlier, sought to transform a rural and tribal society into a modern nation state from the top down during his decade-long reign from 1919 to 1929. Like the United States in post-2001 Afghanistan, Amanullah had a Western-centric approach to modernity, envisioning a state where the material and immaterial aspects of Western society and governance could coexist with

---

Islam.\footnote{Asta Olesen, \textit{Islam and Politics in Afghanistan} (Richmond: Curzon Press, 1995), 118.} A rebellion in Khost in 1924 and a tribal revolt in 1929, however, resulted in Amanullah’s self-exile in Italy until his death in 1960, and the dismantling of the modernisation program.\footnote{Amin Saikal, \textit{Modern Afghanistan, A History of Struggle and Survival} (London and New York: IB Tauris, 2004), 73–92.}

This paper argues that Amanullah’s objective of transforming Afghanistan into a modern nation state through reforms that would centralise authority, unify a fractured society and industrialise the economy, was not only promising, but also achievable. Amanullah mismanaged the implementation of his reforms, however, and alienated most segments of the population. Without an effective bureaucracy to convey and implement his modernisation program, few Afghans understood or benefited from his reforms. Nevertheless, most bore the burden of increased taxation to pay for them. Amanullah also antagonised tribal and religious leaders by curbing their power and prerogatives without ensuring that their interests were protected. Finally, the reforms, which were intended to strengthen the army, effectively crippled and alienated it, leaving Amanullah vulnerable to the widespread resentment that he had inadvertently generated.

The Road to Modernisation

Although modernisation only took centre stage after Amanullah’s ascension to the throne, its seeds had been planted decades earlier during the reign of Habibullah Khan, Amanullah’s father and predecessor who ruled from 1901 to 1919. Departing from the repressive rule from 1800 to 1901 of his own father, Abdur Rahman, the ‘Iron Amir’, Habibullah ushered in a progressive era in Afghanistan through the introduction of reforms, albeit on a limited scale. New roads, factories and Afghanistan’s first hospital and hydro-electric plant were constructed.\footnote{Martin Ewans, \textit{Afghanistan: A Short History of its People and Politics} (New York: HarperCollins, 2002), 85.} Habibullah did not stop at economic reforms, but also established Afghanistan’s first school with a European-style curriculum in 1904, the Habibia College. Furthermore, he allowed the return of political exiles.\footnote{Jeffery J Roberts, \textit{The Origins of Conflict in Afghanistan} (Westport, CT; London: Praeger, date?), 37.} The most prominent of the returning exiles was Mahmoud Tarzi who, during a period abroad, had been influenced by the reformist ideas of the Young Turk movement. On his return, Tarzi established the newspaper, \textit{Siraj al-Akhbar}, to promote his vision of ‘Muslim modernisation’ — that is, a return to the days of the Abbasid Caliphate, where Islam could coexist with a modern and progressive society.\footnote{Vartan Gregorian, ‘Mahmud Tarzi and Saraj-al-Akhbar: Ideology of Nationalism and Modernization in Afghanistan’, \textit{Middle East Journal} 21, No. 3 (1967), 347.}
The new Afghan intelligentsia, formed from students of the Habibia College, drew inspiration from Tarzi’s ideas about Afghan nationalism and modernisation, and by 1909, a reformist movement called the Young Afghans, or mashruba khwahan (constitutionalists), had emerged. Prince Amanullah was involved with the Young Afghans from the beginning. He grew particularly close to Tarzi, not only sharing his vision for Afghanistan’s future, but even marrying one of Tarzi’s daughters, the future Queen Soraya. Amanullah had long been troubled by the state of his country and, even as a child, he was ‘deeply ashamed’ of Afghanistan’s backwardness. With the death of Habibullah in 1919, Amanullah had his chance to rectify this and, with Tarzi as his advisor and foreign minister, he publicly declared that Afghanistan would ‘take its proper place among the civilised powers of the world’.

Centralisation of Authority and Administration

The one objective that no Afghan ruler had truly achieved was unification of the country under a central government, and the establishment of a modern nation state. Thomas Barfield notes that the authority of the central government only applied in its full extent within urban areas, as rural areas remained largely autonomous. The fragile nature of the state ensured that temporary centralisation and unification were dependent on a strong leader, and his death would result in civil war until the next strong leader emerged. Amanullah, however, had an opportunity to break this cycle in 1919 when he waged the Third Anglo-Afghan War and secured Afghanistan’s independence from Britain. With the significant political capital he gained from this success, Amanullah set out to achieve what his predecessors had failed to do, and forge Afghanistan into a modern nation.

In order to centralise administrative functions and political authority, a cabinet and legislature were established for the first time in Afghan history to assist the monarch in governing. In 1923 a de facto constitution, the Basic Codes of the

---

7 Saikal, Modern Afghanistan, 48.
8 ibid.
13 Saikal, Modern Afghanistan, 61.
High State of Afghanistan, was drafted with the assistance of Turkish experts and introduced several new concepts to Afghanistan, such as the freedoms of religion, speech and the press.\textsuperscript{15} The codes also established the rule of law and the concept of a constitutional monarchy, whereby the legitimacy of the king no longer rested on Islam, but on popular legitimacy.\textsuperscript{16} Amanullah also made judicial reforms, including the introduction of new laws based on secular Turkish legal codes, and the codification of *sharia* law, which would be exercised by a new system of secular courts.\textsuperscript{17} Ultimately, the political and legal reforms were intended to transfer power from the religious establishment (*ulema*), which had great influence due to its ability to interpret the *sharia*, to the monarchy and the central government.

Amanullah also sought to bring education under the control of the central government. He believed that the ‘ignorance’ of the Afghan people was responsible for the backwardness of the country. Educational reform was, therefore, ‘closest to [his] heart’.\textsuperscript{18} Prior to 1920 education was controlled by the *ulema* and largely involved rote learning sacred texts and classical works. The national literacy rate was barely two per cent.\textsuperscript{19} A secular curriculum was introduced alongside Islamic subjects, and new schools were established to provide vocational training.\textsuperscript{20} Compulsory education was also introduced in 1924 and, while due to a lack of resources it was not applicable to the entire country, Amanullah did attempt to make more options available.\textsuperscript{21} For instance, new primary and secondary schools were built and night literacy classes were opened for adult learners, at which the King himself occasionally taught.\textsuperscript{22} To emphasise the importance of education, Amanullah made the Medal of Education the highest decoration in the country.\textsuperscript{23}

A centralised nation state also required an army that was loyal to the central government. The army of Amanullah’s predecessors, however, consisted of soldiers nominated by the leadership of the tribes, and they retained their respective tribal loyalties.\textsuperscript{24} Moreover, the army proved to be incapable of defending the nation against foreign threats, as was evident in the near loss to the British during the Third Anglo-Afghan War. Afghanistan may have

\textsuperscript{15} The Basic Codes of the High State of Afghanistan (*Nizam-Namah-ye Asasi-ye Aliyah-ye Afghanistan*) avoided any reference to the term ‘constitution’ to allay fears that it would supplant the *sharia*. See Poullada, *Reform and Rebellion*, 93.

\textsuperscript{16} Olesen, *Islam and Politics*, 121.

\textsuperscript{17} Saikal, *Modern Afghanistan*, 74.


\textsuperscript{20} Poullada, *Reform and Rebellion*, 88.

\textsuperscript{21} ibid., 89.

\textsuperscript{22} Saikal, *Modern Afghanistan*, 75.


\textsuperscript{24} Poullada, *Reform and Rebellion*, 115.
been defeated without the assistance of the Wazir and Mahsud tribes.\textsuperscript{25} In an attempt to rectify this, Amanullah followed the proposals of Tarzi and Turkish military advisors to create a professional army consisting of men from the entire population who were loyal to the central government instead of the tribes.\textsuperscript{26} Universal conscription was introduced for all men of the age of 21 for a period of two years; this was later extended to three years.\textsuperscript{27} As most of the soldiers in the army were veterans and resistant to change, Amanullah reduced their pay fourfold, to just four rupees a month, in order to force older soldiers out of the army.\textsuperscript{28} To compensate for the lower pay, soldiers were to be given a package of benefits, such as cooked food and uniforms.\textsuperscript{29} Inspired by the British use of air power during the Third Anglo-Afghan War, Amanullah also established an air force with assistance from the Soviets, who provided 13 airplanes as well as pilot training.\textsuperscript{30}

### Unification of Society

The nation state that Amanullah envisioned could not be achieved by centralising power and authority alone. It also required the unification of the ‘multi-ethnic, multi-lingual and multi-religious society’ that had only been temporarily unified in the past through the use of force.\textsuperscript{31} Amanullah sought to break down social divisions by abolishing slavery and child labour.\textsuperscript{32} Religious freedom was also granted to minority communities, who were encouraged to play an active role in the state. For instance, Amanullah admitted Sikhs and Hindus to military schools to train as army commanders, offering these communities an opportunity for social advancement.\textsuperscript{33} Amanullah also encouraged Western dress in place of traditional outfits, in the hope of making societal divides less visible.\textsuperscript{34}

The most significant and controversial aspect of the social reforms was Amanullah’s move towards the emancipation of women. For the first time in

\textsuperscript{25} ibid., 111.
\textsuperscript{26} ibid., 114.
\textsuperscript{27} Saikal, \textit{Modern Afghanistan}, 78.
\textsuperscript{28} ibid.
\textsuperscript{29} Poullada, \textit{Reform and Rebellion}, 117.
\textsuperscript{32} Ahmad, \textit{Survival of Afghanistan}, 183–84.
\textsuperscript{33} Emadi, \textit{Repression, Resistance and Women}, 60.
\textsuperscript{34} Poullada, \textit{Reform and Rebellion}, 81. There was also a religious aspect to the clothing reforms. During a speech in Egypt, Amanullah chastised the Egyptians for treating the fez as a requirement of Islam and noted that heabolished traditional clothing in Afghanistan because it had ‘no religious character’. See Ahmad, \textit{Survival of Afghanistan}, 185.
Afghan history, and in most parts of the Islamic world, women were granted freedom of choice in marriage, equal rights to inheritance, a minimum age for marriage and legal protection against abuse. Queen Soraya was a key figure in this emancipation process, and she publicly advocated for change in the roles of women as well as women’s rights to education, employment and divorce. The first primary school for girls was opened in 1921, yet another first in Afghan history. To avoid these reforms being viewed as a covert bid at secularisation, great efforts were made to justify them in Islamic terms. For instance, Amanullah argued that monogamy was inherently more Islamic when he tried to discourage the practice of polygamy. Women were also encouraged to unveil and the veil was portrayed as a ‘tribal custom’ rather than a requirement of Islam. In a public speech, Queen Soraya argued that the reforms were intended to allow Afghan women to contribute to society ‘in the manner of the women of early Islam’.

**The Seeds of Industrialisation**

Amanullah’s grand scheme for modernisation required vast sums of money, but foreign aid was no longer a viable option after the Third Anglo-Afghan War. Not only had the British not rescinded their annual financial subsidy but they also dissuaded other world powers from aiding Amanullah. The Soviets promised Amanullah technical and financial assistance to compensate for the British subsidy, but these promises were left largely unfulfilled. To make Afghanistan economically independent, Amanullah introduced economic reforms, which were intended to begin the process of industrialisation. He began by reorganising and rationalising the tax system, including the abolition of arbitrary taxes, such as those collected for ‘oil for the queen’s hair’. He also introduced the first national budget in the history of Afghanistan and the afghani, a new currency, which replaced the largely valueless rupee. In order to finance industrialisation

37 ibid.
42 The Soviets quickly lost interest in Afghanistan with the diminishing British presence in the country. See Roberts, *The Origins of Conflict*, 42.
43 Poullada, *Reform and Rebellion*, 137.
Amanullah set out to improve agricultural productivity. This was achieved through a policy of land reform, where public land was sold to poor peasants at the low price of ten afghanis per jerib (0.5 acres) of irrigated land.\textsuperscript{46} Amanullah also began planning for the infrastructure requirements of an industrialised economy, with the construction of communication networks such as roads, telegraphic and telephone lines and postal services.\textsuperscript{47} The centrepiece of the infrastructure development was the Great North Road, which would finally provide a direct connection between northern and southern Afghanistan through the Hindu Kush.\textsuperscript{48} By the late 1920s, Afghanistan was also connected by air to Tashkent, Tehran and India.\textsuperscript{49} Amanullah purchased equipment to establish a light industrial sector, including industries such as woodworking, textiles and papermaking.\textsuperscript{50} Students were sent abroad to work as apprentices in European and Persian factories and to gain technical skills, which would be used to develop new industries in Afghanistan.\textsuperscript{51}

\section*{Alienating the Populace}

Amanullah’s modernisation program was so comprehensive that it would have been difficult to implement even in a centralised nation state, much less a largely rural and decentralised Afghanistan. In order for reforms to be implemented throughout the country, a centralised bureaucracy, like the ones that Ataturk in Turkey and Reza Shah in Iran relied on for their own modernisation programs, was imperative.\textsuperscript{52} Amanullah, however, had neither the acumen nor the patience to develop such a capability before advancing with his reform agenda. Lacking political experience, Amanullah’s modernisation plan was based entirely on theory that he ‘learnt through books’.\textsuperscript{53} Furthermore, he allowed his embarrassment at Afghanistan’s backwardness to cloud his judgment, turning a long-term reform process into a rushed scramble to show the world that ‘Afghanistan exists on the map’.\textsuperscript{54} Of the 76 decrees issued by Amanullah throughout his reign, covering every aspect of the modernisation program, 57

\begin{footnotesize}
\begin{enumerate}
\item ibid., 75; Poullada, \textit{Reform and Rebellion}, 135.
\item 74.
\item 141.
\item 247.
\item 139–40.
\item 88–89.
\item 21.
\item M Nazif Shahrani, ‘King Aman-Allah of Afghanistan’s Failed Nation-building Project and its Aftermath’, \textit{Iranian Studies} 38, No. 4 (2005), 662.
\end{enumerate}
\end{footnotesize}
were issued before the Khost Rebellion in 1924.55 Had Amanullah heeded his predecessor Abdur Rahman’s warning that reforms should not be introduced ‘in such a hurry as to set the people against their ruler’, the revolt in 1929 might have been averted.56

In his rush to reform Afghanistan, Amanullah did not consider how his measures, especially the ones intended to benefit the people, would be implemented. In reality, most of Amanullah’s reforms had little effect beyond areas under direct control of the central government.57 For instance, the abolition of slavery and forced labour was in all likelihood ignored outside of Kabul and the neighbouring provinces.58 Amanullah’s strict laws to curb extortion by local officials were never enforced and corruption continued to flourish.59 By 1928, Amanullah was practically pleading with Afghans to cooperate with his modernisation program, lamenting that ‘he advises you, implores you and presses you, but beyond that he can do nothing’.60

The lack of an effective bureaucracy also meant that Amanullah had no means of communicating directly with the vast majority of Afghans about the intention and scope of his reforms. With most news on the modernisation program reaching ordinary Afghans in the form of rumours, dangerous misunderstandings ensued. For instance, Pashtuns, who adhered to a strict tribal code (Pashtunwali) that viewed women as property to be ‘protected’, were horrified by supposed government bans on the veiling of women, when no such policy existed.61 When Amanullah planned to replace turbans with hats, because the former required a greater amount of imported cloth, conservative Muslims feared that it was an insidious anti-Islamic ploy to prevent them from pressing their foreheads to the ground during prayers.62 While conservative and rural Afghans would have found many of Amanullah’s reforms unpalatable, the circulation of rumours and exaggerations intensified the distrust and hostility towards Amanullah and his program.

Although most of Amanullah’s reforms generated little benefit for Afghans living outside urban areas, these ordinary Afghans were saddled with the financial costs.63 In his desire to make Afghanistan appear modern as soon as possible, Amanullah authorised costly projects, such as the new capital

55 ibid., 668.
56 Christopher D Dessaso, Toward Development of Afghanistan National Stability: Analyses in Historical, Military and Cultural Contexts (Kansas: School of Advanced Military Studies, United States Army Command and General Staff College, 2010), 23.
57 Saikal, Modern Afghanistan, 80.
58 Overby, Hard Case of Reform, 19; Poullada, Reform and Rebellion, 144–45.
59 Roberts, The Origins of Conflict, 45.
60 Ahmad, Survival of Afghanistan, 207.
61 Barfield, Cultural and Political History, 185; Overby, Hard Case of Reform, 19.
62 Poullada, Reform and Rebellion, 139.
63 Gregorian, Emergence of Modern Afghanistan, 270.
Dar-ul-Aman, which cost ten million rupees, one third of the state’s annual income.64 Leon Poullada lauds Amanullah’s ‘remarkable feat’ in domestically funding the entire modernisation program with little foreign assistance.65 In reality, this was achieved through a significant increase in taxation. Over the ten-year period of Amanullah’s reign, the land tax increased fourfold, while livestock tax increased two to fivefold.66 Ultimately, the primary effect that Amanullah’s reforms had on ordinary Afghans was a heavier tax burden. Not surprisingly, as the Soviet Central Asian newspaper, Pravda Vostoka, noted at the time, this generated such resentment among the peasantry that they later supported the tribal and religious rebellion.67 In fact, it was an attack on Amanullah’s tax collectors by Shinwari tribesmen that set off the 1929 revolt.68

**Alienating Traditional Power**

Amanullah’s reform program not only cost him popularity among the people, it also gained him powerful enemies among tribal leaders and the *ulema*, the traditional centres of power in Afghanistan who stood to lose most from Amanullah’s plan for a centralised nation state. As Wolfgang Zapf notes, modernisation is ‘not a consensual process, but a competition between modernisers, conservatives and bystanders’, and Amanullah ended up on the losing side.69 Previous rulers were fully aware of the power commanded by the traditional forces in the country and either coopted them with financial subsidies and power-sharing, or controlled them through ruthless suppression.70 Amanullah did neither as he rapidly expanded his own power through his reforms. This made a revolt inevitable.

With the central government assuming responsibility for tasks that had previously been performed at a local level, such as conscription and taxation, tribal leaders found themselves losing power and opportunities for corruption.71 In Amanullah’s drive to prioritise spending, he cut the subsidies to the tribes

64 Barfield, *Cultural and Political History*, 183; The new capital included personal comforts such as a racetrack, a movie theatre and a garage for Amanullah’s car. See Tomsen, *The Wars of Afghanistan*, 73.
67 Gregorian, *Emergence of Modern Afghanistan*, 270.
68 ibid., 264.
70 Abdul Rahman waged numerous wars against rebellious tribes over his 20-year reign, with mass enslavement and executions as punishment. On the other hand, Habibullah made significant concessions to the tribal chiefs and allowed them to influence policy-making in order to improve relations. See Rasanayagam, *A Modern History*, 11–15.
71 Barfield, *Cultural and Political History*, 184.
that previous rulers had instituted.\textsuperscript{72} Moreover, it was not only the purses of the tribal chiefs that were attacked, but also their pride. Traditional ranks and titles were abolished in the spirit of egalitarianism and tribal chiefs were targeted by Amanullah’s anti-corruption drive, which even resulted in the imprisonment of his relative, the Durrani tribal chief Sarwar Khan.\textsuperscript{73} This was considered a gross violation of tribal notions of kinship and loyalty.

The relationship between monarch and ulema had long been a contentious one and Amanullah’s inability to recognise the potential threat posed by the ulema to his reign and reforms proved to be his undoing. Although Amanullah began his reign with a considerable amount of goodwill among the ulema his victory against the British, this quickly eroded due to his foreign policy outreach towards Turkey and Iran, which were viewed as anti-Islam due to their secularisation programs, and the atheist Soviet Union.\textsuperscript{74} Like the tribal chiefs, the ulema detested Amanullah’s centralising policies and social reforms, which were viewed as government intrusions into their traditional areas of influence.\textsuperscript{75} For instance, the centralisation and secularisation of the judicial system threatened a significant source of income for religious officials, who were paid to settle disputes and other local legal issues.\textsuperscript{76} The constitution and laws promulgated by Amanullah also posed a threat to the ulema’s influence and power, which were derived from their role in defining and interpreting religious laws.

More importantly, Amanullah failed to exercise the level of control over the ulema that his predecessors had, even though his modernisation program was a greater threat to the power of the ulema than anything done before. Preceding monarchs understood the necessity of forcing the ulema into submission, and did not hesitate from using force to achieve this. Abdul Rahman threatened mullahs who opposed him with either exile or a departure ‘into the next world’, while Habibullah promptly executed mullahs who publicly opposed his visit to India at the invitation of the British.\textsuperscript{77} Instead of coercing or co-opting the ulema into submission earlier in his reign while he still possessed sufficient political capital, Amanullah did nothing until he was confronted by open rebellion.

In his bid to modernise Afghanistan, Amanullah systematically undermined the power and prerogatives of the traditional power holders in the country, without addressing their concerns or making any concessions in return. Filled with resentment, local ulema and tribal leaders forged a marriage of convenience,

\textsuperscript{72} Gregorian, Emergence of Modern Afghanistan, 271–72; Poullada, Reform and Rebellion, 108.
\textsuperscript{73} ibid., 109.
\textsuperscript{74} Amanullah was even once considered as a possible successor to the former Ottoman Caliphate. See Overby, Hard Case of Reform, 13; Gregorian, Emergence of Modern Afghanistan, 261–62.
\textsuperscript{75} Astri Suhrkei and Kaja Borchgrevink, ‘Negotiating Justice Sector Reform in Afghanistan’, Crime, Law and Social Change 51, No. 2 (2008), 216.
\textsuperscript{76} Poullada, Reform and Rebellion, 120.
which first reared its head in 1924, when Mullah Abdullah, aka the Lame Mullah, succeeded in inciting the Mangal and Jadran tribes in Khost to rebel in a conflict that eventually cost 14,000 lives.78 To deal with the crisis, Amanullah called for a national assembly (loya jirgah) of prominent ulema and tribal leaders and, in exchange for their support against the rebellion and its participants, he was forced to make concessions, such as abolishing restrictions on polygamy and restoring the power of religious judges.79

While much attention has been given to the reforms that Amanullah was forced to sacrifice at the loya jirgah of 1924, with Senzil Nawid going as far as to consider it a ‘decisive victory’ for the ulema, there has been less focus on what Amanullah was allowed to retain.80 The bulk of Amanullah’s reforms in the fields of administration, education, taxation and conscription remained intact. While the tribal leaders and ulema would certainly have preferred having no reforms at all, it is clear that they were willing to compromise so long as their interests were safeguarded. Amanullah refused to acknowledge this, however, and after his European tour of 1928, he increased the pace and scope of his reforms.81 Instead of seeking the ulema’s cooperation, Amanullah publicly derided them as ‘corrupt’ and ‘narrow-minded’, ending their state stipends and forcing religious officials to be trained in state-sponsored schools.82 The final straw came during the loya jirgah of 1928, when Amanullah humiliated tribal and religious leaders by forcing them to wear European clothes.83 When a tribal revolt developed in Jalalabad in 1928, it had the full support of the ulema, who labelled Amanullah as an infidel.84 Amanullah tried to pacify his opponents by repealing several contentious social and religious reforms, but it was too late. He abdicated the throne in January 1929.85

**Alienating the Army**

Although Amanullah angered ordinary Afghans and the tribal and religious leaders, he might still have been able to salvage his rule and reforms had he retained support from the most important institution of all, the army. As the

---

79 Barfield, *Cultural and Political History*, 186.
81 Amanullah’s trip to Europe was a clear contribution to his sheer insistence on modernisation, despite facing resistance. He noted that ‘Paris filled me with joy, Berlin astounded me, London caught my imagination, but it was my homeland which ignited fire in me’. See Shahrani, ‘Afghanistan’s Failed Nation-building Project’, 671.
82 ibid.; Barfield, *Cultural and Political History*, 189.
83 Emadi, *Repression, Resistance and Women*, 64.
85 Roberts, *The Origins of Conflict*, 44.
revolts that toppled his rule were largely localised in the Pashtun-occupied areas of the country, an effective and loyal army should have been able to suppress them.\footnote{Shahrani, ‘Afghanistan’s Failed Nation-building Project’, 664.} The army had long been essential in bolstering the rule of Amanullah’s predecessors, and it was through the army’s support that Amanullah had succeeded in seizing the throne after the death of his father.\footnote{Roberts, The Origins of Conflict, 40.} When Amanullah met Ataturk in Turkey, he was warned that the success of his modernisation program depended on the support of a strong army to suppress resistance.\footnote{Rasanayagam, A Modern History, 21.} Instead, Amanullah’s army failed in both the Khost Rebellion of 1924, where only tribal intervention saved the day, and the final rebellion in 1928.\footnote{Barfield, Cultural and Political History, 187.}

While Amanullah did institute reforms that were intended to create an effective and loyal army, his negligence in implementing the reforms ‘functionally dismantled and ultimately destroyed’ the military.\footnote{Dessasso, Development of Afghanistan National Stability, 25.} The decision to purge the army of its veterans in order to build a professional standing force was not followed up with adequate training, resulting in an inexperienced army that was inferior to its predecessor.\footnote{Poullada, Reform and Rebellion, 118.} Improved benefits such as food and housing, which were intended to compensate for pay cuts, fell by the wayside due to administrative delays and corruption.\footnote{ibid., 117.} This resulted in a loss of morale and professionalism, as soldiers had to accept other employment to support themselves.\footnote{Saikal, Modern Afghanistan, 78.} Amanullah’s support for foreign military instructors from Turkey also served to alienate soldiers and officers who prided themselves on a militaristic tribal culture and were insulted at having to learn warfare from foreigners.\footnote{Shahrani, ‘Afghanistan’s Failed Nation-building Project’, 669.} Amanullah’s military reforms caused recruitment problems and mass desertions, with the army shrinking from over 60,000 men during the time of Abdur Rahman to just 11,000 by 1928.\footnote{Ali A Jalali, Rebuilding Afghanistan’s National Army (Carlisle: US Army War College, 2002), 77.} Not only was the small and inexperienced army incapable of dealing with the tribal uprisings, but Amanullah had also alienated the army to such an extent that masses of soldiers simply defected to the rebels.\footnote{Overby, Hard Case of Reform, 18–19.}

Poullada defends Amanullah’s impact on the military, stating that he made a laudable effort in reforming it, but ‘success eluded him’.\footnote{Poullada, Reform and Rebellion, 118.} But the army’s failure during the Khost Rebellion should have signalled to Amanullah that it was imperative to revise his military program. Amanullah did not listen to advice from his minister of war, Nadir Shah, who warned him about the importance of
Ironically, Amanullah’s negligence in maintaining the military came from his belief that he would always be able to muster the tribes to deal with military threats. As he discovered in 1929, this was not necessarily the case.

Conclusion

Many observers have concluded that the current nation-building effort in Afghanistan is doomed to failure, and have even referred to Amanullah’s failed modernisation effort to demonstrate the futility of such a task. While Amanullah’s program was flawed, however, his experiences do not suggest that such a task is impossible. After all, Amanullah’s reign lasted for a decade, despite increasing hostility towards his reforms. Amanullah was able to secure some measure of support from tribal and religious leaders, even in times of crisis, such as the Khost Rebellion. Unfortunately, Amanullah’s firm belief in the necessity and benefits of modernisation blinded him to the offense and economic hardship that his reforms wreaked on the people and the extent of power and benefits that the tribal leaders and ulema stood to lose in a modernised Afghanistan.

More importantly, Amanullah failed to acknowledge that the support of a strong and loyal army was necessary to protect himself from resistance to his reforms. If the modernisation program had been better managed, taking cultural, tribal and religious sensitivities into account, both Amanullah’s rule and his reforms may have continued. Amanullah failed to do this, however, and his reign and his reforms came to an end in 1929.

Bibliography


99 In July 1923, Amanullah told the people, ‘These are the days of the pen, not of the sword ... therefore send your sons to school. Our martial qualities are sufficient, it is education that we lack.’ See Jalali, Rebuilding Afghanistan’s National Army, 86.


Mechanistic Explanation: Some Limits and their Significance

Dominie Dessaix

Abstract

In this essay I discuss mechanistic explanation (ME), which is a type of explanation characteristic of, though not exclusive to, the biological sciences. ME differs radically from the deductive-nomological (DN) model of explanation around which, until recently, the received view of explanation in the philosophy of science was centered. The aim of the essay is firstly to present and explain ME, and compare it to the traditional DN account. I also consider three potential limits of ME that have been discussed in the literature, relating to ME’s scope, generality, and the completeness of the models it generates. While all of these point to real features of ME, I argue that they are not problematic because of the nature of the phenomena ME aims to explain, as well as the nature of scientific explanation more generally. My conclusion is that no limit to ME is as serious as the consequences of not giving this important kind of scientific explanation a central place in our account of how science explains the world.

Explanation being one of the central aims of science is perhaps obvious. It subsumes some of the other aims that are attributed to science (e.g., understanding how the world works, uncovering reality), and is a favourable companion to others (e.g., successful prediction). Far less obvious is what constitutes a good scientific explanation, and whether one account could fit all sciences, or even all the kinds of phenomena in a single science. Mechanistic explanation (ME) is a type of explanation prevalent in, though not exclusive to, the life sciences. In what follows, I outline a general account of ME, provide some examples, and briefly contrast it with the deductive-nomological (DN) account of explanation: what was, until recently, ‘the received view of scientific explanation in philosophy’ (Bechtel & Abrahamsen 2005: 421), despite some serious criticism (e.g., Cummins...
2000; Kim 1962). I then consider three potential limits to ME. Firstly, I argue that the probable restriction of ME’s scope to certain areas of science, or to some kinds of phenomena in those areas, doesn’t threaten the validity of ME where it is found fruitful. Secondly, I look at the objection that MEs, unlike explanations that appeal to scientific laws (i.e. DN), cannot properly generalise over the phenomena they describe. I argue that this is also unproblematic, and may even be construed as a virtue of ME. Thirdly, I look at Nicholson’s (2012) argument, which states that if we think of biological mechanisms as real world entities, the mechanistic explanations that describe them will fail to be complete, because they must always abstract away from the mechanism’s ‘organismic context’. I show firstly that Nicholson’s discussion conflates real world mechanisms with the models that describe them, and then argue that lack of completeness is not a serious objection to ME, or indeed any scientific explanation. I conclude that although we should be mindful of ME’s limited scope, generality and completeness, these features do not pose problems for the use of ME for at least certain phenomena in the life sciences, and perhaps beyond.

It is first necessary to characterise ME. Minimally, a mechanistic explanation involves presenting a model of the mechanism (the *explanans*\(^4\)) responsible for some phenomenon of interest (the *explanandum*\(^5\)). A *mechanism* is an entity or structure, such as an organ (though a mechanism need not be delineated by such clear physical boundaries\(^6\) that performs a function by virtue of its component parts, operations and, importantly, their spatial and temporal *organisation* (Bechtel & Abrahamsen 2005: 421–23). That is, a mechanism does not just consist in its physical parts, but in the ‘organised interplay’ of the activities of those parts (Soom 2012: 656). Examples of *phenomena* (i.e., functions or behaviours performed by mechanisms) include such things as the heart’s pumping blood, metabolism and memory. Explanation is achieved by a mechanistic model ‘rendering a phenomenon intelligible’ by showing *why* the phenomenon happens in the way that it does.

Consider, for example, the function of the heart to pump blood. What constitutes a mechanistic explanation of the heart *qua* blood-pumper are the parts of heart (atria, ventricles, valves, etc.), the operations performed by these parts (e.g., contraction and relaxation by the four chambers), and their organisation (blood flows from each atrium to valve and ultimately into the circulatory system) (Bechtel & Abrahamsen 2005: 424). Spatial memory is also, in principle,

\(^4\) *Lat.* that which does the explaining.
\(^5\) *Lat.* that which is to be explained; i.e., the thing or event that requires explanation.
\(^6\) Craver (2007: 141–44) discusses how delineating the boundaries of a given mechanism must be done with respect to *explanatory relevance*, not just on the basis of compartmental (i.e., physical) boundaries.
amenable to ME (though it doesn’t yet have a completed model). Knowing that the hippocampus is crucially involved in our capacity to represent the spatial layout of our environment means we know that the hippocampus is a crucial component of the mechanism responsible for behaviours related to the processing and memorisation of spatial features of the environment (Soom 2012: 656; Craver 2002: 89).

Now, before considering the possible limits of ME, I shall describe the DN account of explanation, and draw some comparisons between it and ME, in order to demonstrate how the latter is distinctive. Just as in ME, a DN explanation consists in an explanandum (on the DN view, a sentence describing the phenomenon to be explained) and an explanans (the sentences that do the explaining). DN requires the fulfillment of four conditions: the explanans must 1) contain a statement of a law of nature, 2) include an empirical statement amenable to testing and, 3) consist only in true statements; and, the explanandum must 4) follow logically from the explanans (Fetzer 2013). The central thesis is that a phenomenon is explained by subsumption under natural law (Hempel 1966: 50ff). For example, one could derive the position of Mars at some future time (the explanandum) with an argument (the explanans) containing Newton’s laws of motion, the inverse square law and empirical statements about the initial conditions, such as the mass, position and velocity of both Mars and the sun (Woodward 2011). Contrast this with Bechtel & Abrahamsen’s (2005: 422) example of metabolism (cellular respiration), which if made to fit into the DN model of explanation, would use law-like statements like the following:

Under specified conditions in the last phase of cellular respiration, the ratio of oxygen molecules consumed to ATP molecules produced does not exceed 1:3. (Bechtel & Abrahamsen 2005: 422)

Though true, such a statement, and the DN explanation it would form part of, can do little explanatory work. By not detailing the mechanism of metabolism in which oxygen and ATP molecules play a part, such an explanation cannot tell us what lies behind that ratio.7 For these cases, what we want to show is not ‘that a phenomenon fits within the nomic8 nexus’ but rather ‘how a phenomenon fits into the causal nexus’ (Craver 2007: 39 on Salmon 1984). That is, what is interesting and explanatory here is finding out what causes and follows each stage of metabolism, and how all the components and processes relate to one another. We want to know, in short, how cellular respiration works. Law-like generalisations such as those above are not illuminating. Whatever the virtues

---

7 See Bechtel & Abrahamsen (2005) for more information on the ME of cellular respiration.
8 Here ‘nomic’ just means ‘relating to a law’.
of the DN model in some areas of science (or otherwise, for the DN explanation has its own issues\(^9\)), for these sorts of biological processes it doesn’t yield satisfactory explanations.

There are three further features of ME that make it appropriate where it is most notably employed, viz. the life sciences. Firstly, it is inherently multi-level, aiming to explain higher level phenomena by reference to mechanisms at lower levels (Soom 2012: 657). This means that ME has a recursive structure; i.e., a mechanism can be contained in another mechanism, which can be contained in another mechanism, and so on. The heart, for example, is both a mechanism for pumping blood and a component of the larger mechanism of the circulatory system (Bechtel & Abrahamsen 2005: 425). DN explanations, as separate arguments, do not give this sort of unity of explanation, which is so appropriate for biological phenomena. Secondly, ME is flexible with regard to how it is presented: like DN explanations, MEs can be presented linguistically, but they might also employ diagrams, three-dimensional models, computer simulations, and so on. Thirdly, ME is not simply either correct or incorrect, unlike (the deductive part of) a DN explanation must be, and is thus more obviously consistent with actual scientific practice, in the life sciences at least, where mechanistic models are continually filled in and refined (Craver 2007: 113f).

A perspicuous potential limit to ME is its scope: are MEs only applicable to some phenomena in biology? After all, the example of DN discussed above (that of predicting Mars’ position) does look like a clear case of subsumption under law, and not at all like ME terrain. Furthermore, current accounts of ME may not even work for all of biology. For example, McManus (2012) argues that causal-mechanistic explanations are inconsistent with developmental mechanisms. Yet some restriction of ME’s scope to certain fields or phenomenon type is neither worrisome nor surprising. For ME to count as a valid kind of scientific explanation, it is sufficient that it proves fruitful for some phenomena in some sciences. And it is worth bearing in mind that the scope of ME is still an open question: it could turn out that the usefulness of ME extends out from the biological sciences. Ramsey (2008), for example, discusses how organic chemists consider mechanisms explanatory. In any case, the fact that mechanisms are widely and successfully used in explanations in many areas of biology (Brechtel & Abrahamsen 2005: 421f), coupled with the fact that not all phenomena are amenable to ME (e.g., Glennan’s (2002) ‘genuinely singular events’\(^{10}\)), suggests

---

\(^9\) Such as the problem of explanatory irrelevancies: it is possible to create valid DN arguments that patently miss their explanatory target, as in the famous example that, from the law ‘All males who take birth control pills regularly fail to get pregnant’, and fact ‘John Jones is a male who has been taking birth control pills regularly’, we can derive the conclusion ‘John Jones fails to get pregnant’, for which the premises don’t offer a satisfactory explanation. See Woodward 2011 for this and other issues faced by the DN model.

\(^{10}\) Glennan (2002: 349f) gives the example of a person first meeting her partner: such events typically ‘involve a confluence of events that were not to be expected and will not be repeated’, and there isn’t a sufficiently
not that ME is on the wrong track, but that we require a pluralistic account of explanation. This probable limit of scope accepted, I will move on to consider two possible limits to ME that I think invite more discussion. First, I look at the apparent limited generality offered by ME explanations, especially when contrasted with DN explanations, then move on to consider Nicholson’s (2012) argument that an ‘ontic’ view of biological mechanisms (roughly, conceiving of mechanisms as real things in nature) doesn’t sit well with the fact that MEs are always incomplete due to the necessity of abstracting away from the context of the whole organism.

Scientific laws of nature are universally quantified generalisations. Thus explanations appealing to them allow us to generalise to those unobserved cases that the law subsumes. It looks to be a virtue of DN explanations that generalisation is automatic. Mechanistic models, on the other hand, can be highly ‘specific’ and ‘context-bound’ (Bechtel & Abrahamsen 2005: 437). Scientists working on a given ME might take into account a number of local factors that are only applicable to the single individual or type of entity under study. So, one might ask, how can a mechanistic explanation help us account for what we should expect to find in unobserved cases? Bechtel & Abrahamsen (2005: 437) suggest that we appeal to an alternative kind of generalisation inspired by prototype and exemplar theories. The general idea of such theories is that an item’s membership of a certain category can be a matter of degree, depending on how typical the item is of that category. For example, psychological research conducted by Rosch (1975, 1978) showed that we do not treat all bird species as equally good members of the bird category, but instead tend to classify, for example, robins as highly typical members (one of the ‘birdiest’ bird species), and penguins as highly atypical (considerably less birdy). In a parallel way, we could classify mechanisms as belonging to a more general mechanism-type based on those mechanisms’ similarity to one another, and their resemblance to a prototype. This could apply, for example, to the various mechanisms of protein synthesis in different individual organisms and species, and in different cell types of the same species (Bechtel & Abrahamsen 2005: 438).

So, by allowing ourselves a fuzzier conception of membership based on similarity and prototypicality, we may get sufficient generality of mechanism types, without precluding cases with important variations, whether at the level of the individual, species, or cell type. In addition, I suggest that by embracing a kind of functionalism that brings disparate mechanisms under the same mechanism type as long as they perform an analogous function, we might have another generalising tool, of which the appropriateness in the biological

accurate yet general enough mechanistic model of partner-meeting corresponding to individual events of this kind. This leads me to agree with Glennan in thinking that ME can’t explain ‘fragile actual sequences’ (2002: 349f). Of course, laws don’t seem to be operating here either, so DN would hardly serve better.
sciences has independent empirical support. Consider, for example, that left- and right-handed subjects show different patterns of cerebral lateralisation (i.e., use different neural mechanisms) when performing the same behaviour (i.e., to produce the same phenomenon) (Stoyanov et al. 2012: 1; Sun & Walsh 2006: 655). This non-automatic, fuzzier sort of generalisation, which might also appeal to functional roles, should be thought of as a consequence of the appropriateness of ME to the scientific domain, where it is most often used. After all, the biological world is teeming with variation, so we shouldn’t expect to find a single mechanistic explanation to work for countless slightly varied mechanisms of, say, protein synthesis. The limited generality of ME is thus more a virtue of rather than a limit to ME’s explanatory power.

Now I shall consider a final possible limit to ME, that of completeness, as discussed by Nicholson (2012). So far I have been talking about mechanisms both as real things in nature, and as described by models in scientific explanation. Nicholson (2012), however, has argued against what he calls an ‘ontic conception of mechanisms’ (conceiving of the mechanisms described by ME as real entities in the world), at least in the context of biology. He points out that biological mechanisms are never described completely by MEs, which abstract away from the ‘organismic context’ necessary for the proper function of the mechanism (2012: 159). That is, in actual scientific practice, biologists only include in a given ME those features they deem most relevant to bringing about the phenomenon of interest, necessarily presupposing many features of the wider context that make the mechanism work as it does. Referring to Craver’s own normative requirement on ME: that it must fully account for the explanandum behaviour, not merely a subset of its features (Craver 2007: 161), Nicholson then argues that on these terms, MEs are necessarily incomplete (2012: 159).

This leads Nicholson to argue for an ‘epistemic’ view of mechanisms: we should consider the mechanisms described by ME as ‘heuristic models that facilitate the explanation of phenomena’ (2012: 158). For if we think of the mechanisms described by our MEs as just approximating what happens in the natural world, then we should expect that those descriptions will be incomplete. While I ultimately agree with taking a roughly ‘epistemic’ view of ME, Nicholson’s discussion unfortunately conflates the real world mechanisms that ME tries to explain, with the mechanistic models that figure in our explanations. He defines a biological ‘causal mechanism’ as ‘a step-by-step explanation of the mode of operation of a causal process that gives rise to a phenomenon of interest’ (2012: 153). The latter is clearly a definition of a mechanistic explanation, not of a mechanism itself. Nicholson also claims that it is problematic to think of mechanisms as ‘real systems in nature’ or as ‘autonomous complex systems […] which constitute and operate within the organism’ (2012: 159). Yet I take it that mechanisms themselves can and should be seen as real things out there in
the world, which a *mechanistic explanation* tries to capture and describe. As for any kind of scientific explanation, ME aims to describe the components, operations and organisation of real entities in nature, entities whose existence is not dependent on the biologists who study them. Though Nicholson nowhere denies this realist position, his talk of ‘mechanisms’ to mean *mechanistic models* or *explanations* runs together issues of scientific realism with the issue with which he is really concerned, namely, the issue of how *completely* ME capture those real world mechanisms.\footnote{There is a genuine debate in the literature between the ‘ontic’ and ‘epistemic’ conceptions on ME, but it is rather more subtle than Nicholson’s (2012) treatment. Basically, the ontic view says that the mechanism itself being causally responsible for a phenomenon is what explains that phenomenon, whereas the epistemic view of what makes ME explanatory emphasises the role of conveying understanding about how the mechanism produces a phenomenon. See Illari (2013) for discussion and a persuasive argument that these views are not mutually exclusive.}

But as far as the incompleteness of MEs is concerned, Nicholson’s (2012) argument is convincing: we should bear in mind that mechanistic models are just *models*, and MEs just *abstractions* of the phenomena they describe. I stress, however, that such incompleteness is not problematic. That we are presently unable to access and specify all the ways in which the organismic context influences individual mechanisms in minuscule detail, and that it may never be possible to do so, says only that we cannot construct ‘technically complete’ mechanistic models. We can still aim for (and hopefully achieve) high quality sketches, which may even reach a sort of ‘practical’ rather than ‘technical’ completeness, in that they may explain the mechanism fully when certain (very many) background conditions are specified (e.g., for the mechanistic explanations of memory, ‘the brain must be contained in a living individual of species x’). But more important, I think, is the fact that incompleteness is not an impediment to, but a necessary feature of, many kinds of scientific explanation. If we want to give a satisfactory explanation of a wide range of similar, non-identical instances, our explanation must be sufficiently general, and this simply means abstracting away from the details. We could not, for example, give a mechanistic explanation of heart function in vertebrates (diverse creatures with diverse organismic contexts) without such abstraction. A certain ‘incompleteness’, then, is necessary for some scientific explanations to be explanatory, and thus does not amount to a serious objection to ME.

This essay has discussed ME, comparing it to the DN model and describing some of the features that make ME suited to the phenomena it has been used to explain, most notably in biology. I considered potential limits, none of which were found to threaten ME’s validity as a fruitful sort of scientific explanation. If ME is indeed limited to some areas of science or some kinds of phenomena, this points to the need for a pluralistic account of explanation, not to some defect of ME. Also, ME’s inability to automatically generalise, as the appeal to laws allows
the DN model to do, does not make ME explanatorily weaker, but different in the kind of generalisation it can provide, which is argued in turn appropriate to the biological phenomena it describes. Finally, I argued that although Nicholson’s (2012) presentation of his argument confounds real mechanisms and mechanistic models, I agree with the ‘epistemic’ view of ME that he suggests, which accepts that ME are necessarily incomplete. I argued, however, that such incompleteness is not problematic, and is indeed an important feature of much scientific explanation. Thus the limit to our understanding of science, if we don’t come to grips with this crucial kind of scientific explanation, is far graver than any limit to ME itself.

Bibliography


—— (2007), Explaining the Brain, Oxford University Press.


**Other Sources Consulted**


Dependent Yet Defiant: The Implications of Unilateralism in Iraqi Kurdistan

Sebastian Klich

Abstract

Since 2009 the Kurdistan Regional Government (KRG), which is recognised in Iraq’s 2005 constitution as the governing body of the Kurdistan Region of Iraq, has repeatedly signed oil exploration contracts with international oil companies (IOCs) that the Iraqi federal government considers unconstitutional. In light of its ethno-nationalist heritage, the KRG’s unilateral actions have widely been perceived as part of a broader strategy to gain juridical independence. This essay examines the implications of the KRG’s provocative behaviour and assesses the contention that the KRG is driven by underlying motivations to establish an independent state. It is also demonstrated here that interpreting KRG unilateralism as a move towards secession is simplistic; the KRG is economically reliant on Baghdad and, although Turkey has recently fostered closer relations with the Iraqi Kurds, its historical opposition to an independent Kurdish state will continue. Rather, it is argued that the KRG signing contracts with IOCs is a strategic manoeuvre designed to maximise the economic autonomy of the Kurdistan region in the context of an increasingly fragmented Iraq.

Introduction

In 2005 the status of the Kurdish Regional Government (KRG) was officially recognised in the Iraqi constitution, enshrining the right of the Iraqi Kurds to localised power over ‘all administrative requirements of the region’.1 Within the constitution an important distinction was made between the right to manage existing and future oil fields, the former falling to the federal government and the latter to the regions.2 Or so it seemed. Contrasting views and interpretations

2 ibid.
have been contested by constitutional experts worldwide, debating who can legally manage what, according to the constitution, and a clear consensus is yet to emerge. Baghdad insists that all contracts for oil exploration and production must be developed in partnership with the central government. The KRG refutes this position with its actions and, since 2009, it has signed over 50 contracts with international oil companies (IOCs), contracts that Baghdad considers to be illegal. Given the majority of the KRG’s income streams directly from the Iraqi national budget, this aggressive economic behaviour, which is defiant of Baghdad, has stimulated a whirlwind of suppositions through commentary and media worldwide speculating that the time of an independent Kurdistan is drawing near.

In analysing the implications of such provocative behaviour, this essay will firstly consider the contention that the KRG is strategically positioning itself for the creation of an independent Kurdish state, reinforcing the historical trajectory of Kurdish nationalistic aspirations, alongside the KRG-directed cultural change and economic growth that has recently taken place. The essay then examines the consequences of the KRG’s oil-sector management in the broader domestic context of Iraq. The KRG’s domestic political motivations outside of the independence debate are examined by highlighting the relationships of multifaceted tensions that exist between Erbil (capital of the KRG) and Baghdad, the widening of sectarian rifts, and the attempted secession of other Iraqi provinces. Next, neighbourly cooperation with Turkey is scrutinised by assessing the motivations that underpin this purposeful collaboration of regional convenience, and shining a revealing light on the implications it has for Kurdish independence.

After outlining the sturdy barriers to a separate Kurdish state being established, the essay finally draws the analysis together to argue that the current obstacles to achieving independence are too great for this to be the intent of the Iraqi

---


5 This essay uses the terms ‘Iraqi Kurdistan’ and ‘The Kurdistan Region of Iraq’ interchangeably, referring in both instances to the region controlled by the KRG. The title of ‘The Republic of Kurdistan’ refers to the hypothetical state that would likely be established if the Kurdistan Region of Iraq were to attain independent statehood and no longer exist as a federal region. The ‘central government’ refers to the Iraqi government in Baghdad.
Kurds, and considers that the KRG signing contracts with IOCs is a strategic ploy for maximising Kurdish economic autonomy in the context of an increasingly fractured Iraq.  

The Republic of Kurdistan?

Speculation on the prospect that the KRG is seeking to establish an independent state has been influenced by the vitality and virility of the Kurdish national movement. The Kurds like to think the ‘earthly paradise’ that exists in the north of Iraq today is the cumulative result of a century of struggle and hardship. The Kurds were the primary beneficiaries of the 2003 Iraq War. The perceived benefits came, however, at a high price. A full historical account of the struggle of the Kurds is beyond the scope of this essay, but the following points are key to appreciating the strong and resilient sense of Kurdish nationalism.

The Kurdish people can be traced as an identifiable ethnic group for more than 2000 years, though it was at the turn of the twentieth century that they established a sense of national community. The end of the Ottoman Empire and World War I held out the promise of Kurdish independence in the ill-fated 1920 Treaty of Sevres. Rejection by Turkey and continuing war led to the 1923 Treaty of Lausanne, a political compromise that pacified Turkey, but sacrificed a Kurdish state. After Iraq gained its independence in October 1932, Sheikh Ahmed Barzani led the first (failed) uprising demanding Kurdish autonomy. In 1946 the Kurdish Democratic Party (KDP) was established, announcing the independent Mahabad Republic of the Kurds, which lasted until 1947 when Iran’s forces invaded.

The 1970 peace accord was the first recognition of basic Kurdish autonomy by the Iraqi government. The period 1988–91 saw Saddam Hussein’s genocidal ‘Anfal’ campaign attempt the extermination of the Kurds and resulted in the killing of between 50,000 and 100,000 people. In 1992 the first elections of the Kurdistan region were held, and the KRG was established. Since the KRG’s

---

6 The focus of this essay is the implications of KRG unilateralism. As such, an examination of the Kurdish nationalist movements in neighboring countries is beyond its scope.
7 The Economist, ‘Northern Iraq’.
11 ibid., 132.
13 McDowall, A Modern History of the Kurds, 179.
14 ibid., 296–97.
16 ibid., 359.
formation, a territorially based nationalism has developed in Iraqi Kurdistan, which takes precedence over the ethnically oriented form, distinguishing the Iraqi-Kurd movement from its regional ethnic counterparts. The strength of this movement was reflected in an informal referendum held across the Kurdistan region simultaneously with national elections of 2005; 98 per cent of voters expressed their preference for independent statehood. This brief historical background highlights the fortitude and proven endurance of the Kurdish nationalist movement, laying a foundation for the claim that the KRG’s recent economic defiance of Baghdad is the natural trajectory of an ongoing aspiration, demonstrably sustained over time, to form an independent state of Kurdistan.

The KRG are gradually enforcing cultural and linguistic changes in Iraq’s Kurdish society. This reinforces perceptions of seeking a more distinctive and separatist identity as the possible forerunner of independence. Schools in the Kurdish region now teach English and Turkish from the first grade. Arabic and Kurdish are the official languages of Iraq, as enshrined in the constitution. The KRG, however, enforced a change in the curriculum in Iraqi Kurdish schools in 2006 that saw the teaching of Arabic, which dominates the rest of the country, being delayed until the fourth grade. This was before the influx of foreign direct investment began to reshape the economic landscape of Kurdistan, countering arguments that claim the shift in language teaching is a response to the emerging Western currents flowing throughout the commercial boom. The decision to prioritise English and Turkish over Arabic suggests that the KRG has attempted to steer the cultural development of its region in a new direction, away from the crumbling control of centralist Iraqi institutions in Baghdad.

The debate surrounding the prospective independence of Kurdistan has become increasingly topical because of the region’s significant economic growth. Before relations between Erbil and Baghdad deteriorated into the recent verbal and military hostilities, the KRG benefitted from a lucrative agreement in which the central government paid for the production costs of oil companies in Kurdistan. This was in addition to the 17 per cent of the annual Iraqi budget already being paid to the KRG, equating to 95 per cent of revenue. The Kurdistan region

---

22 ibid.
also chose to implement liberal investment laws that have attracted a wealth of foreign direct investment (FDI). The success of this approach was displayed in 2011 when *FDI Magazine* ranked Erbil as the fifth most promising destination in the Middle East for potential FDI. The combination of FDI and the lucrative oil management agreement with Baghdad fuelled unprecedented economic growth in the Kurdistan region, improving the Kurdish lifestyle significantly and setting it apart from the rest of Iraq, whose citizens can only envy relative luxuries such as near 24-hour electricity. The economic development of the Iraqi Kurds projects an image of prosperity, intensifying the cultural separation from the rest of Iraq. In this context, defying the central government by signing contracts that Baghdad declared unconstitutional can be perceived as a renewed and defiant push for independence.

Historically, Kurdish nationalism has proven to be steadfast in the face of continuous persecution. Culturally, schools have implemented curriculum changes that are steering Kurdish children’s learning away from the Arab heritage of Iraq. Economically, the Kurdistan region has flourished, and the KRG’s investment laws have helped to motivate and manoeuvre the Kurds to make rapid inroads in parts of their economy. Structurally, the Kurdistan Region of Iraq has a government that is considered legitimate by the people. It has entered into relations with other states and, apart from the border disputes in the south (which are not uncommon between states, especially in the Middle East), it exists within a defined territory. As Stansfield points out, these characteristics align precisely with the definition of a state in the 1933 Montevideo Convention on the Rights and Duties of States. Cumulatively, these historical, cultural, economic and structural features of the Kurdistan region portray a people and government in search of a state. In this context, the KRG signing contracts with IOCs could conceivably be an attempt to establish the income stream necessary to rejuvenate, promote and maintain a drive for independence. To accurately discern the KRG’s motivations, however, a more comprehensive examination is needed.

**A Push for Independence or Disregard for the Iraqi Central Government?**

In order to objectively examine the strategic motives of the KRG, the wider domestic context must be taken into consideration. As the thriving oil sector

---

26 ibid.
of the Kurdish region has been a major driver of its economic prosperity, the dispute over the constitutional rights to manage the region’s natural resources can be perceived as the predominant source of hostility between Erbil and Baghdad. In reality, it is the tip of the iceberg.

Although tensions have existed over natural resource management for some time, the downward spiral of relations began after the current Iraqi Prime Minister Nouri al-Maliki reneged on the conditions of power sharing that were reached in the Erbil Agreement in November 2010. The Kurds played a vital role in the negotiation of the Erbil Agreement, which secured the formation of the second Maliki government on the condition that the Prime Minister would cede control of key security ministries of the state. In response to Maliki’s disloyal and manipulative politicking, Kurdish President Massoud Barzani and Iraqi President Jalal Talabani, the leader of the Patriotic Union of Kurdistan (PUK), attempted to orchestrate a vote of no confidence to remove Maliki from power. This failed attempt only fuelled Maliki’s increasingly dictator-like leadership.

The disputes over territory, including the oil-rich city of Kirkuk, have been another propellant, or even accelerant, of this political blaze. In the months following the attempted vote of no confidence, the Kurdish forces known as the Peshmerga, and the Iraqi Security Forces (ISF) were deployed against each other in the province of Kirkuk. The political skirmish that grew out of Maliki’s refusal to relinquish power, combined with the ongoing dispute over territory, shows that hostility between Erbil and Baghdad runs much deeper than disagreements over natural resource management. This suggests that the KRG’s obstinate decision to continue signing IOC contracts could in fact be an antagonistic and aggressive move in a wider political battle, separate from any intent to establish an independent Kurdistan.

Displays of assertive economic policy by the Kurds are aggravating existing sectarian rifts in the delicate political landscape of Iraq. Since the withdrawal of US troops at the end of 2011, large segments of the Sunni Arab population in Iraq have regarded the actions of Maliki as deliberately marginalising and

---

31 Although Iraqi Kurds are predominantly Sunni, KRG cooperation with Sunni and Shi’a factions is driven by political expediency, not sectarian belief. This is unlikely to change given that the disputed territories to which the KRG lay claim are populated largely by Sunni Arabs, while the central government is predominantly
subordinating the Sunni Arab minority.\textsuperscript{32} The latest wave of sectarian difficulty rose in 2011 when Maliki issued an arrest warrant for the Sunni Vice-President Tariq al-Hashemi.\textsuperscript{33} In contrast with the views of Anthony Blinken, who in 2012 described Iraq’s ‘progress towards a more normal political existence’ as being ‘remarkable’,\textsuperscript{34} sectarian violence has since worsened. In late 2012 and early 2013, demonstrations erupted across Iraq in Fallujah, Ramadi, Tikrit and Mosul.\textsuperscript{35} Many of these protests called for the Iraqi government to step down and for the police force to withdraw, accusing Maliki of being a dictator.\textsuperscript{36} A dramatic increase in sectarian-motivated bombings has since ensued.\textsuperscript{37} The last time Maliki faced such growing unrest throughout the provinces of Iraq was during the height of the civil conflict in 2006–07, when he was allied with the KRG and also had the support of the US troops.\textsuperscript{38} US troops have since departed and relations between Erbil and Baghdad have soured dramatically. By signing contracts with IOCs, the KRG is defying Maliki’s predominantly Shi’a government, strengthening the claims of much of the Sunni Arab population that Maliki is an illegitimate leader.

The KRG’s management of its natural resources is fuelling more than just sectarian violence; it is acting as an achievable template and encouraging wider decentralisation of Iraq’s energy sector and demands for greater regional autonomy. Kurdistan, Iraq’s only federal region, has become the safest and arguably most prosperous part of Iraq. Other provinces perceive this prosperity to be a result of the KRG’s increasingly independent natural resource sector and are attempting to gain a similar level of autonomy. Richard Wadsworth, the CEO of Canadian firm Sonoro Energy, recently claimed to have signed an asphalt licence agreement with the province of Salahuddin.\textsuperscript{39} Similarly, a preliminary prospectus for Canadian company Oryx Petroleum states that an agreement with the province of Wasit has been signed.\textsuperscript{40} The heterogeneous Ninewa province,
which is made up of Christians, Sunni Arabs and Kurds, is experiencing a similar phenomenon, with the local Christian population calling for its own autonomous area within Ninewa.\textsuperscript{41} This is the same disputed territory that includes the Bashiqa oil exploration block issued to ExxonMobil by the KRG.\textsuperscript{42}

These expressions of provincial independence coincide with burgeoning demands by Sunni Arabs for their own regional federalism.\textsuperscript{43} The perceived autonomy of the Kurdistan region, which the KRG flaunts by signing contracts with IOCs, is fuelling demands for similar levels of independence throughout the rest of the country and encouraging other provinces to manage Iraq's local energy resources for their own regional benefits. These destabilising factors are chipping away at what hopes remain, and what international agencies and organisations might wish, for Iraqi unity.

A Regional Relationship of Convenience

An increasingly intensified collaboration between the Turkish Government and the KRG appears to be integral to furthering the economic autonomy of the Iraqi Kurds. The foundations of this partnership have been laid over the past five years as Turkish businesses have invested heavily in the growing economy of Iraqi Kurdistan.\textsuperscript{44} This has evolved into a thriving trade relationship, with US$12 billion in KRG–Turkey trade taking place in 2011.\textsuperscript{45} In 2010 Turkish trade with the KRG was greater than Turkish trade with Syria, Lebanon and Jordan combined.\textsuperscript{46} This relationship recently became central to the KRG's economic future, as Genel Energy has begun building a pipeline between the Taq Taq oilfield and an existing Iraq–Turkey pipeline.\textsuperscript{47} Although an official agreement between Erbil and Ankara is yet to be confirmed, Turkish Prime Minister Recep Tayyip Erdogan openly stated in March 2013 that he perceives any such agreement to be within the limits of the Iraqi constitution.\textsuperscript{48} This indicates Erdogan is resolved to persist with such an agreement, reinforced by his recent proposal for Turkey to create an escrow account through which it

\begin{itemize}
  \item \textsuperscript{41} Denise Natali, ‘The Politics of Kurdish Crude’, \textit{Middle East Policy} 19:1 (Spring 2012) 112.
  \item ibid.
  \item ibid.
  \item ibid.
  \item Orhan Coskun, ‘Iraqi Kurdistan Poised to Pipe Oil to World via Turkey’, \textit{Reuters} \url{http://www.reuters.com/article/2013/04/17/iraq-kurdistan-oil-idUSL5N0D310920130417} (17 April 2013) accessed 19 April 2013.
\end{itemize}
could manage the distribution of the oil revenues for both Baghdad and Erbil.\footnote{Denise Natali, ‘Can Turkey Leverage Kurdish Crude?’, \textit{Al-Monitor} \url{http://www.al-monitor.com/pulse/originals/2013/04/turkey-kurdish-oil-crude-baghdad-leverage-energy-policy.html?utm_source=&utm_medium=email&utm_campaign=6850} (8 April 2013) accessed 19 April 2013.}

On the surface, the forging of closer relations between Erbil and Ankara appears to be a possible gateway for the level of economic and regional autonomy that could support an independent Kurdish state.

A closer examination of the motivations underlying improved relations between the KRG and Turkey is required before it can be considered an economic platform from which an independent Kurdish state could potentially be launched. A major contributing factor to the newfound friendship between Turkey and the Iraqi Kurds is the conflict in Syria. In the current civil war, two main Syrian Kurdish opposition parties have come to prominence in the north of the country. The Democratic Union Party (PYD) is a Syrian Kurdish branch of the Kurdistan Workers’ Party (PKK), the rebel Kurdish group, widely considered to be a terrorist organisation, which is now based out of Iraqi Kurdistan and fights for Kurdish recognition within Turkey. The other main organisation is the Kurdish National Council (KNC), which is comprised of 12 smaller Kurdish parties that combined to form the KNC under the guidance of KRG President Barzani in late 2011.\footnote{Stansfield, ‘The Unraveling of the Post-First World War State System’, 280.} Barzani has used his leverage over the KNC to negotiate an agreement between it and the PYD that aimed to achieve joint leadership within Syrian cities previously dominated by the PYD.\footnote{Ben Gittleson, ‘Syria’s Kurds look to Iraqi Minority for Support’, \textit{New York Times} \url{http://www.nytimes.com/2013/01/31/world/middleeast/31iht-m31-kurds.html?_r=0} (31 January 2013) accessed 10 March 2013.} Ankara considers the PYD to be a duplicate of the PKK, which fought against Turkey for Turk–Kurdish autonomy in a bloody conflict that started in 1990.\footnote{Kemal Kirişçi, ‘Turkey’s Kurdish Challenge’, in \textit{The Kurdish Policy Imperative}, eds Robert Lowe and Gareth Stansfield (London: Royal Institute of International Affairs, 2010) 59.} Hence, Ankara is cautious of the rising power of the PYD in Syria, perceiving it to be a potential springboard for the PKK.\footnote{International Crisis Group, ‘Syria’s Kurds: A Struggle Within a Struggle’, \textit{Middle East Report} 136 (22 January 2013) accessed 10 April 2013.} Barzani’s negotiations were, therefore, received warmly by Ankara as it attempted to dilute the power of the PYD within Syria.\footnote{Stansfield, ‘The Unraveling of the Post-First World War State System’, 280.} This suggests that by improving relations with the KRG and supporting its interpretation of the Iraqi constitution, Ankara is strategically positioning itself to gain greater influence over the KRG in the hope of minimising the perceived threat posed by Syria’s PYD. Ankara’s foreign policy objectives run deeper still.

Turkey’s alliance with the KRG will also help to check any aspirations that the KRG may have of establishing a wider Kurdistan. Superficially, Barzani’s attempt to foster power sharing in his neighbouring region is unsurprising. KDP–PUK
infighting in the 1990s restrained the Iraqi Kurdish nationalist movement. Barzani’s facilitation of a compromise between the PYD and KNC can, therefore, be seen as the President spreading the wisdom from his hard-learned lesson. The ethnic Kurdish homeland, however, extends across southern Turkey, northern Iraq, eastern Iran and north-western Syria. Turkey has violently oppressed the PKK for over two decades, sending a clear message about its stance on the prospect of Kurdish autonomy outside of Iraq. Although Ankara and the PKK have outlined a peace arrangement, its implementation has been problematic, troubled by underlying distrust. Similarly, in relation to Barzani’s negotiations with the Syrian Kurdish factions, not enough time has passed to quell Turkish suspicions that the president of the KRG may seek to rule over a wider Kurdish population that extends beyond the borders of Iraq. This same bordering ally, which may be perceived as a potentially willing partner in furthering the economic autonomy necessary to establish an independent Kurdistan, has a long and bloody history of opposing any transgressions or incursions of its own claimed territories. This has been the underlying motivation for Turkey’s historical and ongoing opposition to an independent Kurdish state, a steadfast position that reveals the political expedience driving Turkey’s recent cooperation with the KRG.

The Realities of Independence

Economic dependence on Baghdad firmly limits the ability of the KRG to establish its independence. The central government currently supplies Erbil with 17 per cent of the federal Iraqi budget, equating to US$10 billion annually, or about 95 per cent of its revenue. This income stream has been crucial to the socioeconomic development of the region. The KRG’s attempts to increase its economic autonomy portray a hope for economic independence in the future. The chief executive of Genel Energy, the IOC pumping oil from the Kurdish fields of Taq Taq and Tawke, claims that the production capacity of the Kurdish region by the end of 2013 could reach close to one million barrels per day. This would be a sizeable increase of the current capacity of 400,000 barrels per day, a figure that is yet to be achieved because of the continuing dispute with Baghdad.

Although the prospect of considerable growth in Kurdistan’s oil sector appears promising, the IOCs that it is assumed will be extracting the oil are the same IOCs that are owed vast sums of money. Under the agreement described earlier, the central government is responsible for paying these companies. Baghdad, however, in deeming these contracts unconstitutional, has largely refused to pay. This casts doubt over any figures predicting future increases in Kurdish oil exports, as such an increase is reliant on the continued commitment of IOCs that are currently owed an accumulative total of US$4 billion. With relations between Erbil and Baghdad disintegrating, this dispute is unlikely to be settled, leaving the KRG reliant on its portion of the national Iraqi budget, thereby limiting any strategy to seek independence.

The city of Kirkuk and the surrounding disputed territories are integral to the economic independence needed to fund an independent state. One of the controversial elements of the KRG’s contracts with IOCs relates to some of these locations. For instance, three of the six blocks contracted to the American firm ExxonMobil are located in disputed border areas in the provinces of Kirkuk and Ninewa. Although the central government has refused to pay the production costs of these ventures, it has not tried to physically intervene or use military force to prevent exploration going ahead. The armed hostilities that took place towards the end of last year between the Peshmerga and the ISF, however, suggest that if the KRG tried to secede, it could only include the disputed territories in its borders, if it was able to take them by force. Although the scale of the conflict is difficult to predict, Baghdad’s military power is mightier than that of Erbil, as it is equipped with US-supplied F16s and M1 tanks.

Furthermore, Turkey may be fostering closer economic ties with the KRG, but it has also made clear that it does not support the Kurdish claims for the disputed city of Kirkuk. If the oil in Kirkuk and the surrounding disputed territories is unavailable to the KRG, then it is predicted that it could not export enough oil to replace its current share of the Iraqi national budget. The centrality of the disputed territories to the economic autonomy of the KRG

---

61 ibid.
62 Natali, ‘Who Will Pay IOCs in Iraqi Kurdistan?’.
67 Alaaldin, ‘Now is Not the Time for an Independent Kurdistan’.
is a definitive barrier blocking any alleged march to freedom, persuasively challenging the notion that signing contracts with IOCs is a strategic ploy directly aimed at achieving independence.

International relations with Iraqi Kurdistan have been rapidly increasing, but this does not translate into tangible or practicable international support for an independent Kurdistan. As previously stated, ‘the capacity to enter into relations with other states’ is a key characteristic of the 1930 Montevideo convention. The KRG has a proven capacity for inter-state relations, but there remains a strong and formidable opponent of Kurdish secession: the United States. Although the United States worked closely with the Kurds during the invasion, occupation and reconstruction of the country, it has maintained an uncompromising stance on pursuing its ‘One Iraq’ policy. When hostilities surged in the disputed territories during late 2012 in what threatened to be the first military engagement between the Peshmerga and ISF in the post-Saddam era, Washington contacted Barzani directly to make clear the strong US opposition to the prospect of secession.

The increasing regional division between Sunni and Shi’a dominated states has reinforced Washington’s view. Given Ankara’s souring relations with Baghdad and Tehran, Washington is concerned that an independent Kurdistan, in its current close relations with Turkey, could drive Baghdad even further into the embrace of Tehran, strengthening Iran’s ‘arc of influence’. The staunch US resistance to Kurdish independence will influence its allies and, most likely, prevent them from supporting secession within Iraq, potentially undermining the recognised legitimacy of an independent Kurdistan.

To contend that the KRG’s signing of contracts with IOCs is geared towards obtaining independence is to assume that a clear strategy has been devised in Erbil. In reality, disagreement between the major parties and political players in the domestic politics of Iraqi Kurdistan raises serious doubts as to whether a clear or a unified direction exists. Since the official agreement between the KDP and PUK in 2007, both parties have put their differences behind them and cooperate for the stability of their region. Differences of opinion still arise however.

During the political row with Baghdad over budget allocations earlier this year, the PUK and KDP disagreed over how best to react to the situation. After boycotting the vote to approve the national budget, which was then passed in their absence, the possibility of withdrawing from the parliament in Baghdad

---

70 Alaldin, ‘Now is Not the Time for an Independent Kurdistan’.
was considered by members of the KDP.\(^{72}\) The PUK quickly rejected this call, claiming it would cause greater difficulties for the KRG.\(^{73}\) Withdrawing ministers from the Baghdad parliament is not as extreme as claiming independence. The difference of opinions between the PUK and KDP on how best to manage relations with Baghdad regarding the fallout over the national budget, however, suggests that the KRG is lacking a coherent or consistent strategic plan.

**Conclusion**

Striving for greater economic autonomy by continuing to sign contracts with IOCs has incited a wave of sanguine speculation about the possible or planned future of the Kurds within Iraq. However great the number of journalists and academics to jump on this bandwagon of optimistic predictions, and whatever the extent of the speculation, the fact remains that crystal-ball gazing and analysis of covert intentions is a perilous art. The reality of Turkish foreign policy, the continued economic dependence on Baghdad, the strong opposition of Washington and the warfare that would likely ensue if the Kurds tried to claim Kirkuk, together form a formidable obstacle to Kurdish statehood being realised in the near future. Furthermore, the disunity and infighting within the Kurdish camp casts doubt on the prospect of a grand plan paving the road to the establishment of a self-determining state.

Independence may be currently out of reach, but the limits of Kurdish autonomy are yet to be clearly demarcated, and notional or conceptual boundaries can be pushed by commercial opportunities. The KRG’s contracts with IOCs are a strategic manoeuvre designed to maximise the economic autonomy of the Kurdistan region. Exacerbating domestic sectarian rifts, reinforcing the regional Sunni–Shi’a divide, and encouraging attempts at secession by other oil-producing provinces are ultimately destabilising results of this ambitious, but ultimately, self-interested gambit, which viewed through the lens of powerful players outside the region, can well be considered dangerous and threatening to any remaining hopes for a secure and unified Iraq. While dreams of independence remain constrained, the message from Erbil is loud and clear: even without an independent state, Iraqi Kurdistan and its resources belong to the Kurds.


Bibliography


Natali, Denise. The Kurdish Quasi-state (Syracuse: Syracuse, 2010).

Natali, Denise. The Kurds and the State: Evolving National Identity in Iraq, Turkey, and Iran (Syracuse University Press, 2005).


Classification of Arid & Semi-Arid Areas: A Case Study in Western Australia

Guy Leech

Abstract

Arid and semi-arid environments present challenges for ecosystems and the human activities within them. Climate classifications are one method of understanding the climate of these regions, giving regional-scale insight into parameters such as temperature and precipitation. This essay qualitatively compares the application of four different climate classification schemes (Köppen-Trewartha, Guetter-Kutzbach, De Martonne and Erinç) to a transect across Western Australia. The classification schemes show a distinct climatic gradient from the wet coast to the dry continental interior, with the most pronounced gradient located near the coast. Strong agreement is found between weather stations and classifications for wet and dry years, indicating that the weather systems on the coast also govern weather inland. Historically wet and dry periods, as well as long-term drying trends were also identified for the coastal regions, which could affect the livelihoods of many people. A long-term drying trend is also evident at Kalgoorlie-Boulder, the most arid, inland town. These results show that climate classifications can identify major trends and shifts in climate on a decadal or yearly basis, and it is hoped that they will become more widely used for this form of analysis.

Introduction

Arid and semi-arid environments make up a large portion of the Earth’s surface (Figure 1), and present challenges for human ecosystems located within them. These regions are generally defined as having low average rainfall, often associated with high temperatures, which impose fundamental limits on animal and plant populations, and on human activities such as agriculture (CSIRO 2011, Ludwig & Asseng 2006, Ribot et al. 2005, John et al. 2005, Vörösmarty et al. 2000, Watson et al. 1997). The challenges presented by arid and semi-arid
environments may be amplified in response to global climate change, as rainfall is predicted to both increase in variability and decline overall in many of these regions (Charles et al. 2010, Lioubimtseva et al. 2005, Ma et al. 2005, Indian Ocean Climate Initiative (IOCI) 2002, Watson et al. 1997). These changes are expected to particularly affect settlements and ecosystems on the boundaries of arid or semi-arid environments (Fitzpatrick et al. 2008, Lioubimtseva et al. 2005, Ma et al. 2005, Camill & Clark 2000, Kepner et al. 2000, Watson et al. 1997), where an increase in aridity (i.e., the spatial extension of the arid zone) could have consequences ranging from biodiversity loss associated with ecosystem collapse, to crop failures and famine (Parmesan 2006, Ribot et al. 2005, Walther et al. 2002, McCarty 2001, Watson et al. 1997). The potential impacts of increasing aridity make identifying the location and extent of arid and semi-arid areas, and determining if trends exist in their climate, important areas of research.

Various methods have been developed to provide insight into the spatial extent of arid and semi-arid areas, of which the most widely known and used are climate classification schemes. These schemes range from those employing biophysical factors, such as precipitation, temperature and vegetation type, to classifications based on the movement of distinct atmospheric air masses or incoming solar radiation, through to schemes that classify climate by regional vegetation and soil types (Sturman & Tapper 2006, Trewartha & Horn 1980). The climate of a region is then classified based on the relationship between some or all of these factors, from which regional characteristics (e.g., vegetation) can then be derived. It is important to note that human worldviews are evident in all classification schemes, either as the basis used to delineate areas appropriate to different human needs (e.g., agriculture) or in the naming of classifications based on human experience (e.g., ‘frequent-fog’) (Sturman & Tapper 2006, Stern et al. 2000, Trewartha & Horn 1980).

The Köppen scheme, upon which the Köppen-Trewartha scheme is based, is a widely used climate classification that combines annual and monthly mean precipitation and temperature to classify areas as arid or otherwise (Sturman & Tapper 2006, Trewartha & Horn 1980, Köppen 1936; Baker et al. 2010, Gao & Giorgi 2008). The Köppen-Trewartha scheme modifies these classifications by considering the seasonality of precipitation, and its ratio to temperatures (Trewartha & Horn 1980, Feng et al. 2012, de Castro et al. 2007). There have been modifications made to the Köppen and Köppen-Trewartha schemes to accommodate variations in regional climates and the preferences of different researchers (Guetter & Kutzbach 1990). Other climate classifications include De Martonne (1926) and Erinç (1965), both of which are based on the ratio of precipitation to temperature. Climate classification schemes also exist based on other parameters, such as potential evapotranspiration (Thornthwaite 1948), or the radiation balance (Budyko 1958).
Figure 1: Global map showing arid and semi-arid areas

Source: Millennium Ecosystem Assessment 2005.
This essay investigates these four climate classification schemes to illuminate
differences and similarities between them, and to determine whether any one
provides a more accurate or useful representation of the climate of Western
Australia. An evaluation of the applicability and usefulness of yearly climate
classifications to studying inter-annual variability and long-term trends is
another central aspect of this study. A transect of Western Australia, from the
cost to the arid interior, is used to examine how the climate of different areas is
classified and whether spatial and temporal variation in classification between
schemes exists. Trends in classifications and climate are investigated in relation
to other studies of the region.

Data and Methods

Western Australia occupies a large portion of the Australian continent, with a
correspondingly wide range of climates and biomes; however, the majority of
the state is dry and hot, with high variability in rainfall (Trewin 2006). Data
from a series of weather stations that form a line running inland from Perth, on
the western coast, to the town of Kalgoorlie, in the dry interior, were used as the
basis for this study (Figure 2). The location and transect were chosen because of
the pronounced rainfall gradient from the coast to the arid interior (Charles et
al. 2010), allowing for a variety of possible climate classifications (e.g. see Figure
3). The station names, numbers, coordinates and elevations are listed in Table 1.
These weather stations were chosen to permit analysis of the climatic gradient
from the coast to the interior (Pitman et al. 2004), as well as for their relatively
long and reliable data records. The annual mean temperature and precipitation
totals were obtained from each of these stations for the years 1951–2012, the
longest period for which data from all stations was available from the Bureau
of Meteorology (BoM). The data for Merredin and Merredin Research Station,
Southern Cross and Southern Cross Airfield, and Cunderdin and Cunderdin
Airfield were combined to give data series of suitable length.

Table 1: Information on the weather stations used in this study, from the
coast (top of table) inland

<table>
<thead>
<tr>
<th>Name</th>
<th>Abbreviation</th>
<th>Number</th>
<th>Location (° S; ° E)</th>
<th>Elevation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perth Airport</td>
<td>PA</td>
<td>009021</td>
<td>31.93; 115.98</td>
<td>15</td>
</tr>
<tr>
<td>Northam</td>
<td>NOR</td>
<td>010111</td>
<td>31.65; 116.66</td>
<td>170</td>
</tr>
<tr>
<td>Cunderdin</td>
<td>CCA</td>
<td>010035</td>
<td>31.65; 117.23</td>
<td>220</td>
</tr>
<tr>
<td>Cunderdin Airfield</td>
<td>CCA</td>
<td>010286</td>
<td>31.62; 117.22</td>
<td>217</td>
</tr>
<tr>
<td>Kellerberrin</td>
<td>KEL</td>
<td>010073</td>
<td>31.62; 117.72</td>
<td>250</td>
</tr>
<tr>
<td>Merredin</td>
<td>MMRS</td>
<td>010092</td>
<td>31.48; 118.28</td>
<td>315</td>
</tr>
</tbody>
</table>
Annual mean temperature and precipitation totals were used to classify the climate at each station according to four different climate classifications: the Köppen-Trewartha scheme (KT); a modified Köppen-Trewartha scheme by Guetter & Kutzbach (GK); an extended De Martonne scheme (DM); and the Erinç scheme (ER). The KT scheme compares temperature with precipitation and seasonality of precipitation in order to classify environments as deserts (arid) or steppes (semi-arid) (Trewartha & Horn 1980). Other classifications under this scheme are based on temperatures and seasonality of temperature only (Trewartha & Horn 1980), and are not calculated here but assumed to fall in the subtropical classification as they occur during wetter years (BoM 2006, Stern et al. 2000). The GK scheme is a modification of the KT scheme and changes some of the boundary conditions between climatic zones, including some arid zone classifications, to better account for changes in climate as identified through changes in vegetation and soil (Guetter & Kutzbach 1990). The DM scheme, a ratio of precipitation to temperature (De Martonne 1926), was extended by Khalili (1992) to more accurately reflect desert environments. Finally, the ER scheme, a different ratio of precipitation to temperature, is also used (Kutiel & Türkes 2005, Erinç 1965). The exact definitions used for the classifications under each scheme are given in tables 2–5.

A standard precipitation index (SPI) was also calculated for each year, as an indication of how far a particular year differed from the long-term average rainfall. In this study the long-term average was drawn from all years for which data were available. The SPI is calculated as:

\[
(\ln(x) - u) / o
\]

where \( x \) is the annual precipitation total, \( u \) is the long-term mean and \( o \) is the variance in the long-term mean (Zhang et al. 2009, McKee et al. 1993). The SPI classifications and probabilities of each classification are given in Table 6.
Figure 2: Map of south-west Western Australia, showing the locations of weather stations; those used are in red markers

Source: Author’s own work, using Google Maps.

**Tables of climate classification definitions**

**Table 2:** Definitions of climate used in this study according to the Köppen-Trewartha scheme. \( r = \) annual total precipitation; \( t = \) annual mean temperature

<table>
<thead>
<tr>
<th>Classification</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steppe</td>
<td>70% or more of annual precipitation falls in the summer half of the year, and ( r &lt; 20t + 280 ); or 70% or more of annual precipitation falls in the winter half of the year, and ( r &lt; 20t ); or neither half of the year has 70% or more of annual precipitation, and ( r &lt; 20t + 140 ).</td>
</tr>
<tr>
<td>Desert</td>
<td>As above, but ( r &lt; 50% ) the amount for classification as a steppe</td>
</tr>
</tbody>
</table>

Source: After Trewartha & Horn (1980).
Table 3: Definitions of climate used in this study according to the Guetter-Kutzbach scheme. \( r \) = annual total precipitation; \( t \) = annual mean temperature; \( p \) = percentage of precipitation falling in the coolest six months of the year

<table>
<thead>
<tr>
<th>Classification</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steppe</td>
<td>( r &lt; 2.4t - 0.46p + 55 )</td>
</tr>
<tr>
<td>Desert</td>
<td>As above, but ( r &lt; 50% ) the amount for classification as a steppe</td>
</tr>
</tbody>
</table>


Table 4: Definitions of climate used in this study according to the modified De Martonne scheme. \( AI \) is the aridity index, defined as \( AI = \frac{r}{t + 10} \), where \( r \) = annual total precipitation; \( t \) = annual mean temperature

<table>
<thead>
<tr>
<th>Classification</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per-humid B</td>
<td>( AI &gt; 55 )</td>
</tr>
<tr>
<td>Per-humid A</td>
<td>( 35 &lt; AI &lt; 55 )</td>
</tr>
<tr>
<td>Humid</td>
<td>( 28 &lt; AI &lt; 35 )</td>
</tr>
<tr>
<td>Subhumid</td>
<td>( 24 &lt; AI &lt; 38 )</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>( 20 &lt; AI &lt; 24 )</td>
</tr>
<tr>
<td>Semi-arid</td>
<td>( 10 &lt; AI &lt; 20 )</td>
</tr>
<tr>
<td>Arid</td>
<td>( 5 &lt; AI &lt; 10 )</td>
</tr>
<tr>
<td>Extra arid</td>
<td>( AI &lt; 5 )</td>
</tr>
</tbody>
</table>

Source: After de Martonne (1926).

Table 5: Definitions of climate used in this study according to the modified Erinç scheme. \( AI \) is the aridity index, defined as \( AI = \frac{r}{t} \), where \( r \) = annual total precipitation; \( t \) = annual mean temperature

<table>
<thead>
<tr>
<th>Classification</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Per-humid</td>
<td>( AI &gt; 55 )</td>
</tr>
<tr>
<td>Humid</td>
<td>( 40 &lt; AI &lt; 55 )</td>
</tr>
<tr>
<td>Semi-humid</td>
<td>( 23 &lt; AI &lt; 40 )</td>
</tr>
<tr>
<td>Semi-arid</td>
<td>( 15 &lt; AI &lt; 23 )</td>
</tr>
<tr>
<td>Arid</td>
<td>( 8 &lt; AI &lt; 15 )</td>
</tr>
<tr>
<td>Severe arid</td>
<td>( AI &lt; 8 )</td>
</tr>
</tbody>
</table>

Source: After Erinç (1965).
Table 6: Categories, definitions and probability distributions for the standard precipitation index (SPI)

<table>
<thead>
<tr>
<th>Category</th>
<th>SPI</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extreme (ex.) wet</td>
<td>≥ 2.00</td>
<td>2.3</td>
</tr>
<tr>
<td>Severe wet</td>
<td>1.50 – 1.99</td>
<td>4.4</td>
</tr>
<tr>
<td>Moderate (mod.) wet</td>
<td>1.00 – 1.49</td>
<td>9.2</td>
</tr>
<tr>
<td>Normal</td>
<td>-0.99 – 0.99</td>
<td>68.2</td>
</tr>
<tr>
<td>Moderate (mod.) drought</td>
<td>-1.00 – -1.49</td>
<td>9.2</td>
</tr>
<tr>
<td>Severe drought</td>
<td>-1.50 – -1.99</td>
<td>4.4</td>
</tr>
<tr>
<td>Extreme (ex.) drought</td>
<td>≤ -2.00</td>
<td>2.3</td>
</tr>
</tbody>
</table>


The different classification schemes were qualitatively compared by observing whether the different schemes provided different classifications for the same station, or showed different trends through time. A similar analysis was conducted across the series of stations, to see if spatial trends existed between schemes. Where changes in classification occurred the SPI was used as an indicator of whether the change was a single event or part of an ongoing trend.

Results

Köppen-Trewartha (KT) classifications

The KT classifications for the weather stations are given in Figure 3. The scheme shows a progression from subtropical to desert classifications going inland from the coast, indicating increased aridity. Spatial and temporal coherence of particularly wet or dry years is observed; many stations were classified as a wetter steppe class during a particular year. The known rainfall variability in the area is also displayed, however, with the same station being classified as desert and then as subtropical during the following year (e.g., KBA 1991–92). Long-term drying trends can be seen at CCA, KEL, and MMRS, while a recent drying trend is evident at KBA. A wet period is observed across most stations during 1958–74.
Figure 3: KT classification for selected weather stations. Blank entries indicate insufficient data. Temperate and sub-tropical classifications are assumed, not calculated.

Source: Author’s own work.
Guetter-Kutzbach (GK) classifications

The yearly GK classifications for each weather station are shown in Figure 4. The climatic progression from the wet coast to the dry interior is again observed. Spatial and temporal coherence of wet and dry years is seen, although dramatic changes in classification at the same station between some years are also observed, indicating the high variability of rainfall. A recent drying trend appears at NOR, while a dry period from 1969–79 is observed at most stations.

De Martonne (DM) classifications

The classifications given by the DM scheme are shown in Figure 5. This scheme shows a clear climatic gradient from the wet coast to the dry interior, although there is often a distinct (two or more classes) difference between PA and NOR climates. Again, spatial and temporal coherence are observed in wet and dry years. A long-term drying trend can be seen for PA, while KBA shows a more recent drying trend; a dry period is observed from 1969–80 at KEL, MMRS and SCSCA. High rainfall variability between years is again observed for some locations and years.

Erinç (ER) classifications

The climate classifications under the ER scheme are shown in Figure 6. A gradient from the wet coast to the dry interior is visible, although the change in classification from PA to NOR is often dramatic (two or more classes). Spatial and temporal coherence of dry and wet years are observed, although large differences in classification occur between some years, indicating high rainfall variability. A long-term drying trend is evident at PA and KBA, while a wetting trend can be seen at SCSCA. A wet period from 1963–68 is observed at many stations, while a dry period from 1969–80 is observed at most.
Figure 4: GK classification for selected weather stations. Blank entries indicate insufficient data. Temperate and sub-tropical classifications are assumed, not calculated.

Source: Author’s own work.
Figure 5: DM classification for selected weather stations. Blank entries indicate insufficient data

Source: Author’s own work.
Figure 6: ER classification for selected weather stations. Blank entries indicate insufficient data

Source: Author’s own work.
Standard Precipitation Index (SPI)

The SPI of each station for all years is shown in Figure 7. High annual variability in rainfall is visible, with conditions changing from extreme to normal between many years, and from severe drought to severely wet. A wet period during the 1960s is visible, as is a dry period from 1969–80. A dry period at PA can be seen from 2001–10, which may be an indication of long-term drying. SCSCA appears to show a long-term wetting trend.

Discussion

The climatic gradient from the WA coast, at Perth, to the arid interior, at Kalgoorlie-Boulder, is evident in each of the classification schemes used. The DM and ER schemes illustrate this, and also showed a distinct jump from the climate of Perth (PA) to the next station inland (NOR), showing the influence of the ocean and elevation (NOR is ~155 m above PA [see Table 1]) on climate. The advantages displayed by these schemes are due to the increased number of classification types, which allows greater insight into subtle differences and changes in the climate of an area.

All of the classification schemes examined showed spatial and temporal coherence in wet and/or dry years; i.e., if one station was particularly wet, it was likely that others were as well. This indicates that the climate in this area of Western Australia is governed by the same weather systems, although rainfall decreases as the systems move inland. All the stations, however, also showed dramatic differences in classification between some years, indicating the onset of particularly wet or dry conditions, which reflects the high rainfall variability of the region (Trewin 2006). No scheme showed better insight than another into these phenomena, although the greater number of classifications within the DM and ER schemes was of benefit in making distinctions between moderately dry years and extremely dry years.
The identification of particular wet and dry periods was similar across the four schemes. Each scheme showed a dry period beginning in 1969, though the end date differed. In the KT and GT schemes the dry period ended in 1979, while in the DM and ER schemes it ended in 1980. A wet period ranging from the
early 1960s until 1969 was also identified in the KT and ER schemes. These differences are a likely consequence of the precise yet somewhat arbitrary definitions of the climatic boundaries, rather than actual reflections of the aridity of the area at that time.

Temporal trends in climate differed most between schemes. The KT, DM and ER schemes showed a drying trend at KBA with the broad agreement between schemes indicating that this trend is likely to be a real phenomenon, and should be investigated more closely. The KT and GK schemes showed evidence of a drying trend at NOR, while the ER and DM schemes showed a recent drying trend at PA. The fact that both of these pairs of schemes are based on the same principles (see Data and Methods, above) means that these trends are more likely to be spurious. The latter trend, however, is also supported by the SPI, suggesting that it is a real phenomenon. The ER scheme and SPI also suggest a long-term wetting trend at SCSCA, indicating another area that is ripe for further research.

The drying trends described above corroborate earlier studies of south-west Western Australia that identify decreasing rainfall in the region from the late 1960s (Hennessy et al. 2007, Power et al. 2005, IOCI 2002, Smith et al. 2000, Hennessy et al. 1999). The identification in this study of a particularly wet period in the 1960s, followed by a distinctly dry decade, reveals, however, that the drying trend observed in previous studies may be exaggerated, as the 1970s appear drier when compared to the preceding decade. The use of individual stations in this study, as opposed to the areal averages generally used in previous studies, provides greater insight into exactly where these drying trends are most evident, although the results reported here conflict with some previous small-scale work. For example, Pitman et al. (2004) predicted that decreases in rainfall should be expected near the coast, with increases in rainfall further inland; the results presented here suggest a coastal drying trend, but also suggest a drying trend inland at KBA, and a wetting trend at SCSCA, which is also inland.

In terms of methodology, this study shows the usefulness and applicability of yearly climate classifications for studying long-term trends. The ability to detect long-term and emerging trends is of vital importance in climatic studies; the fact that climate classifications, generally assumed to be static and requiring long periods of data, are able to provide this insight on a decadal (and sometimes yearly) basis is an important result. It is hoped that the use of classifications for studying climate can be continued, and that the clear and accessible nature of these classifications may aid communication of climate and climate change concepts.

The limitations of the results and analysis presented here primarily stem from the limitations of climatic classifications in general. The use of discrete cut-offs, especially when combined with a small number of weather stations, can lead to
Classification of Arid & Semi-Arid Areas: A Case Study in Western Australia

arbitrary divisions between areas, although this is offset somewhat in the DM and ER schemes because of their many classifications. Additionally, the classification schemes used here only make use of temperature and precipitation data; other schemes that are based on different parameters would be of value for further comparison. Some data and statistical limitations are also present in this study: the calculation of SPI from all available years of data means it may downplay recent trends, and the lack of data for some years, together with the combining of data from different stations, introduces potential errors in the analysis.

Future research in this area should concentrate on the particular drying or wetting trends identified above, or on comparing the results above to those obtained from other classification schemes. Further data collection or reanalysis of weather stations across Australia would benefit future studies of this nature.

Conclusion

This essay examines differences between four climate classification schemes, using data from a transect of weather stations in Western Australia. A distinct climatic gradient is identified from the wet coast to the dry continental interior, and it is most pronounced at the coast. Strong agreement is found between weather stations and classifications for wet and dry years, indicating that the same weather systems govern rainfall in this region. Historically wet and dry periods are also identified. Long-term drying trends are found for the coastal regions and the most arid, inland town; these trends could affect the livelihoods of many people. The use of climate classifications is shown to be useful for identifying long and short-term climatic trends, although they are less useful for understanding spatial variation. It is hoped that future studies on climatic trends may consider the use of these, or other, climate classifications.

Acknowledgements

I would like to thank Janette Lindesay for guidance on this project.

Bibliography


Indian Ocean Climate Initiative (IOCI) 2002, Climate Variability and Change in Southwest Western Australia, Indian Ocean Climate Initiative Panel, Perth.


Framing Gender Inequality: Millennium Development Goal 3 and the Post-2015 Agenda

William Lutwyche

I will personally ensure that the fight for the empowerment of women is at the heart of the international process I am co-chairing to renew the Millennium Development Goals.

UK Prime Minister David Cameron, July 2012

Abstract

Millennium Development Goal 3 (MDG3) has become the universal barometer by which global gender equality progress is judged. This essay argues that the framing of MDG3 and its corresponding indicators of progress obfuscate the complexities of gender equality. Whilst the reductionist nature of MDG3 has allowed for increased awareness of the issue, it de-contextualises and obfuscates the complex nature of gender inequality. This argument is illustrated through tracing the use of the three MDG3 indicators of progress: (i) to eliminate gender disparity in educational enrolment; (ii) increase the share of women in non-agricultural wage employment; and, (iii) the proportion of seats held by women in national parliaments. In response to the criticisms of oversimplified framing of MDG3, this essay analyses how including a consideration of gender-based violence and a twin-track approach in the post-2015 era could, at least in part, address these issues. The current reframing of gender equality in the post-2015 era fails to directly address these matters. Broadly, the analysis in this essay fundamentally brings into question the legitimacy of the progress proclaimed as part of MDG3.

Introduction

Since the Millennium Development Goals (MDGs) were adopted in 2000, there has been an unprecedented level of public, political and financial support for gender equality and empowerment in developing nations. As the 2015 target
date for achieving the MDGs approaches, a number of processes are underway within the United Nations (UN) system and its member states, academia, policymakers and civil society to formulate a new global development agenda. For gender and development, this proposed global framework represents another chance — perhaps the only one — to secure the future of gender equality in development discourse and practice. In this context it is necessary to ask, has gender inequality and empowerment been sufficiently represented in the current framework?

Whilst gender is inherent in other MDGs, this essay is limited to examining MDG3, which endeavours to ‘promote gender equality and empower women’. This essay argues that the framing of MDG3 and its corresponding indicators of progress obfuscate the complexities of gender equality and development practice. While the reductionist nature of MDG3 has allowed for increased awareness, it de-contextualises and obfuscates the complex nature of gender inequality. This conclusion fundamentally questions the legitimacy of progress proclaimed under MDG3. The essay is divided into three sections. Firstly, a brief background and context to the debate is given by introducing the MDGs and MDG3. The second section examines each indicator of progress underlying MDG3 and how gender equality and empowerment are framed to measure progress through gender disparity in educational enrolment, non-agricultural wage employment and the percentage of women in national parliaments. The third section critiques gender-based violence as an essential variable missing from MDG3. It is evaluated and considered for inclusion in a post-2015 framework. Following this, the essay examines the recently proposed macro-level twin-track approach to achieving gender equality. This strategy envisions a reinvigorated stand-alone goal on gender equality in conjunction with gender being mainstreamed across all new post-2015 objectives. These points of analysis clearly demonstrate that the MDG framework is an inadequate reflection of global progress on gender inequality. A gendered post-2015 framework should encompass improved sex disaggregated data collection and redefined progress indicators that can be applied and evaluated universally.

Background

In September 2000, leaders from 189 countries, including 147 heads of state and government, gathered at the UN and unanimously adopted the Millennium Declaration — a commitment to a future based on peace, prosperity and justice (UN 2000). To achieve this, countries endorsed the MDGs, a set of eight goals aimed at development and poverty reduction, with quantitative targets to be achieved by 2015. The goals are based on income deprivation, hunger, gender equality, education, the environment, health and global partnerships for
development. They unequivocally present the alleviation of global poverty as an ethical imperative for all states and international organisations (Hulme & Fukuda-Parr 2009). By 2001 the MDG framework of eight goals included 18 targets and 48 progress indicators. The MDGs are now commonly accepted as the universal framework for measuring development progress globally, nationally and regionally (Clemens, Kenny & Moss 2007; Fukuda-Parr, Greenstein & Stewart 2013; UN 2005: 10). Accordingly, development progress associated with gender equality is measured in line with the indicators underpinning MDG3.

It is universally agreed that the world is a long way from achieving gender equality. The MDGs, however, have been cited as a ‘powerful stimulus’ for raising the profile and importance of alleviating gender inequality globally (Organisation for Economic Co-operation and Development (OECD) 2013). Scaling up investment and action on gender equality has the dual advantage of addressing widespread inequality and accelerating development progress overall (United Nations Development Fund for Women (UNIFEM) 2012). Given its importance, how MDG3 and the forthcoming post-2015 agenda frame gender equality is critical to ensuring development effectiveness and sustainable change. Kabeer (2005), however, questions the meaning of gender equality and empowerment, and how it is framed through a goal-oriented strategy. Additionally, questions have been raised regarding the sufficiency of using only one educational goal and four indicators to measure progress on global gender equality (Elson & Balakrishnan 2012). The next section of this essay unpacks this analysis and shows that MDG3 is framed to de-contextualise gender inequality in order to incite mobilisation and awareness. This approach comes at the cost of obfuscating the reality of development, bringing into question any progress cited under MDG3.

The Framing of MDG3

The MDGs have been useful in drawing attention to pervasive gender-based deprivation in the developing world and the need for a coordinated global effort to address this prevalent issue. The framing mechanisms used to measure MDG3 progress have, however, been extensively criticised. As stated previously, MDG3’s progress is framed through one target: ‘to eliminate gender disparity in primary and secondary education, preferably by 2005, and at all level of education no later than 2015.’ This target was defined as the ratio of girls’ to boys’ enrolment in primary, secondary and tertiary education. Three other indicators were added to MDG3: the ratio of literate females to males among 15–25 year olds, the share of women in non-agricultural wage employment and the proportion of seats held by women in national parliaments. This section determines the extent to which educational enrolment and its successive indicators sufficiently
represent global progress on gender equality and empowerment. It is important to first define these terms and establish a criterion against which the indicators of progress can be measured.

**Gender Inequality and Empowerment**

The UN (2012) defines gender inequality in terms of capabilities (education, health and nutrition) and opportunities (economic and decision-making). Similarly, the World Bank (2011) defines gender inequality through equality under the law, equality of opportunity and equality of voice. Grown et al. (2003) combine these two approaches and suggest two domains for measuring the equality of women in relation to men. Firstly, the ‘capabilities’ domain refers to basic human abilities as measured through education, health and nutrition. This domain reflects the MDG3 indicators of gender disparity in educational enrolment rates. Secondly, the ‘access to resources and opportunities’ domain refers to the opportunity to use or apply the aforementioned capabilities through access to economic assets and resources (Grown et al. 2003: 3). This domain reflects the third indicator of women in wage employment in the non-agricultural sector. Women’s empowerment is closely related to, but goes beyond, gender equality, encompassing not just women’s condition relative to men’s, but also their ‘agency’ and power to make choices (Smee & Woodroffe 2013).

Malhorta et al. (2002) and Kabeer (2005) assert that empowerment should be measured through the ability to make choices. Put simply, Kabeer (2001, cited in Malhorta, et al. 2002: 6) defines it as, ‘The expansion in people’s ability to make strategic life choices in a context where this ability was previously denied to them.’ This definition is useful as it contains two elements that distinguish empowerment from other related concepts, such as equality: firstly, the idea of process and change from a condition of disempowerment and, secondly, that of human agency and choice. Kabeer (2005: 15) argues that the empowerment of women in MDG3 should be centred on forms of ‘transformative agency’. These do not simply address inequality through quick fix mechanisms; they challenge power relations and provoke sustainable changes to the structures of patriarchy. Overall, when framing indicators to measure gender equality and empowerment, there must be equal capability and access to opportunity and resources in conjunction with transformative agency and choice. Through this definition, a critique of each indicator illustrates that MDG3 is a narrow representation and measurement of gender equality and empowerment progress. The way in which MDG3 is framed fails to recognise diversity of experience and communicates progress as if all ascriptions are universally applicable. This sense of de-contextualisation de-legitimises proclamations of progress.
Eliminate Gender Disparity in Educational Enrolment

Extensive research highlights the strong links between education, gender equality and empowerment outcomes. For example, education increases the likelihood that women will look after their own well-being and with that of their family. A study in Zimbabwe found education to be the primary factor that increased the likelihood of women accessing contraception and antenatal care, which improves maternal survival and well being (Kabeer 2003). Further, education can directly challenge and change power relations within the household. In Sierra Leone, educated women were found to have more leverage in bargaining in their household and to take more control over spending domestic income than uneducated women (Kabeer 2003). In this regard, increased enrolment is linked with better access to resources. It is wrong to assume, however, that an increase in enrolment for females will immediately correlate with the benefits such as those outlined. Despite its potential, how MDG3 frames the current education goal is narrow in scope and obfuscates contextual change. This can be illustrated through how it is measured.

Care must be used in interpreting measures of educational equality because perfect equality (1:1) may indicate equality of deprivation rather than equality of opportunity. An example is the case of Ethiopia, where there is only a 3 per cent difference between male and female secondary education enrolment figures. While there is close to parity between the sexes in secondary enrolment, the overall rate of enrolment is only 37.5 per cent (World Bank 2013). This indicates a problem of widespread educational deprivation, rather than an issue of equal access to opportunities. Further, studies suggest changes associated with education are conditioned by the context in which it is provided and the social relationships that it embodies and promotes (Kabeer 2005). For example, where women’s role in society is defined purely in reproductive terms, education is seen in terms of equipping girls to be better wives and mothers, or increasing their chances of getting a suitable husband. These are legitimate aspirations, given the realities of society but they do not equip them with a sense of transformative agency that empowers them to question their subordinate status (Kabeer 2005). All these factors affect the transformative potential of education systems around the world. Despite these contextual variables, Sen and Mukherjee (2013) argue that gender equality in primary and secondary education has begun to epitomise the notion of gender equality and empowerment. Such a limited understanding represents limited progress from ways of thinking witnessed in previous decades. The importance of agency and choice, and recognising the shift in power relations between men and women, has become less important. The framing of MDG3 to measure gender equality purely through educational outcomes has caused it to revert back to merely meaning equality of achievement in meeting certain basic needs (Sen & Mukherjee 2013).
The limitations to education as a route to empowerment and equality do not negate the positive findings nor dismiss their inclusion in a gendered post-2015 framework. While parity in access to schooling is the first step towards gender equality in education, it should not be assumed that its effects would be uniform across all contexts, nor should it be interpreted as a true reflection of development progress. The framing of global gender equality progress should move beyond simple quantitative data to include indicators such as rate of school completion. Despite being hard to measure, qualitative-based data, such as the quality of education and content of the curriculum, should be considered for a post-2015 framework. Such indicators are more likely to reflect challenges against existing patriarchal structures, thus better tracking overall progress on gender equality and empowerment. The level of educational attainment is also strongly linked to entry into the labour market.

Increase the Share of Women in Non-Agricultural Wage Employment

The rise of non-traditional agricultural export production in Africa and Latin America has led to a rise in wage employment for women in large-scale production units. Studies suggest that this income has induced a number of economic improvements for women and their families, including better control over how their money is spent (Kabeer 2005). For example, in Honduras, women working in maquiladoras (assembling manufactured goods for export) earned higher wages than workers elsewhere, and reported improvements in household relationships and help in domestic work from males. They were more likely to have voted in elections and to feel they carried some weight with the government (Hoyos et al. 2008). This case study reflects the potential for progress through access to opportunities and resources. The framing of this indicator of progress, however, also comes with widespread scepticism.

Firstly, this indicator only measures the presence or absence of work, and not the quality of work itself (Anker 2002 cited in Grown et al. 2005). Women’s access to paid work can give them a greater sense of self-reliance, greater purchasing power and overall choice, but if it is undertaken in conditions that exploit their labour or health, its costs may outweigh its empowerment potential. Further, if the division of labour in domestic chores and child care is not renegotiated, such employment can become disempowering, as it becomes a ‘double burden’ in addition to women’s pre-existing responsibilities. The International Labour Organization’s (ILO) Decent Work agenda provides a framework for promoting access to employment and equal treatment. Indicators in this framework include gender-disaggregated unemployment rates, the female to male earnings ratio, and occupational segregation by sex. This comprehensive equality measurement should be considered for a post-2015 gender framework. Such an approach
better highlights the contextual nature of gender equality. Framing gender equality progress through the increased participation of women in national parliaments has the potential for empowerment through transformative agency.

**Participation in National Parliaments**

As women represent half of the world’s population, they should be entitled to equal representation in national parliaments. Such an achievement represents the most ambitious of the three forms of change singled out to measure progress on women’s empowerment. Yet progress has been slow, with the percentage of parliamentary seats held by women increasing from 11.6 per cent in 1995 to only 19.7 per cent by January 2012 (UN 2012). This indicator is, however, correlated with various positive outcomes. For example, a study by Dollar et al. (2001 cited in Grown et al. 2005) found the quality of governance and the level of corruption fell in many countries as women’s representation in national parliament increased. While this goal has great potential, the way in which it is framed de-politicises and simplifies the process in which women are moving into national parliaments.

At present, women who enter national parliaments are not generally drawn from the ranks of the poor, nor is there any guarantee that they are more responsive than men in parliament to the needs of poor women (Kabeer 2012, p. 21). Further, it is an imperfect proxy for tracking aggregate levels of female empowerment because it fails to examine whether women have power in national parliament to make decisions or choices. An increased female presence across governance structures clearly has the transformative potential to change unjust practices.

But if the women are drawn from a narrow elite or have no constituency to which they must answer, their representation could be interpreted as token. Measuring the percentage of women in national parliaments should be accompanied by, and indeed built on, greater participation in local governance processes in a post-2015 framework. Evidence from India suggests that an increase of women in *panchayats* (local governance councils) has made women more likely to take advantage of state services and demand their rights, and also improved the provision of services for men (Chattopadhyay & Duflo 2004). Further qualitative data, which measure female agency and influence to effect policy and outcomes, should be included.

This section argues that the framing of development progress under the heading of MDG3 de-contextualises and obfuscates the complex nature of gender equality. Fundamentally, this finding brings into question the legitimacy of progress proclaimed as part of the MDG banner. For example, an increase in women’s access to education may improve their chance of a good marriage or their capacity to sign their name on a contract or document, but unless it
provides them with the analytical capacity and courage to challenge unjust practices, its potential for transformative change is limited. The MDGs fail to highlight such important variables and real people are represented by abstract concepts and statistical figures to be moved up and down in the charts of progress. Given their potential for transformative change, however, these targets and indicators should not be abandoned when the MDGs expire in 2015. Rather, their definitions should be reconsidered and measurements improved to find a balance between the value of simplifying development for awareness purposes and recognising the complexities of development practice. Elements currently missing from MDG3 and newly proposed approaches to gender equality in the post-2015 era have been proposed to address this pervasive issue. The framing of such approaches is analysed in the next section.

Reframing a Gendered Post-2015 Framework

Gender relations, like all social relations, are multi-stranded: they embody ideas, values, and identities; they allocate labour between different tasks, activities, and domains; they determine the distribution of resources; and they assign authority, agency, and decision-making power. This means that gender inequalities are multi-dimensional and cannot be reduced to some single and universally agreed set of priorities (Kabeer 2005: 22).

The year 2015 will be a time to evaluate the targets and indicators of progress articulated by the MDGs. It also presents as a chance to assess the adequacy and efficacy of MDG3. Reflecting upon Kabeer’s (2005) statement above, a universally agreed upon set of priorities or strategies that address the issue of gender equality and empowerment is unlikely to be found. How progress on gender equality is reframed in the post-2015 era should seek to address the issues of legitimacy and oversimplification for which MDG3 can be criticised. In response, various new approaches have been proposed. This section first examines the inclusion of gender-based violence, an element said to be missing from MDG3, in a post-2015 framework. Secondly, it analyses the validity of framing gender equality through the proposed twin-track approach, which intends to mainstream gender across all post-2015 goals and be accompanied by another stand-alone goal.

Gender-Based Violence

It is estimated that gender-based violence is as serious a cause of death and incapacity among reproductive-age women as cancer; it is a more common cause of ill health among women than traffic accidents and malaria combined (Grown
et al. 2005). It is recognised as a development constraint and its consequences are strongly linked to the desired outcomes of MDG3. For example, it affects children's schooling. A study in Nicaragua found that 63 per cent of children of female victims of violence had to repeat a school year and left school an average four years earlier than other children (Larraín et al. 1997 cited in Grown et al. 2005). Secondly, it has widespread economic multiplier effects. These include increased absenteeism, decreased labour market participation, reduced productivity, lower earnings and savings and lower intergenerational productivity (Grown et al. 2005). For example, it is estimated that in Chile domestic violence reduced women's earnings by $1.56 billion in 1996, or more than 2 per cent of GDP (Morrison & Biehl 1999). Violence against women restricts their ability to use their capabilities and take advantage of opportunities, thereby reinforcing gender inequality. Its inclusion in a post-2015 framework is questioned, however, because of the lack of reliable data and the difficulties associated with quantifying progress.

Accurate and reliable statistical data on gender-based violence is infrequent because of under-reporting by victims and under-recording by authorities (Grown et al. 2005). Thus, the collection of comparable data is rare. The state in many instances sees it as being beyond its scope of influence, as violence typically occurs within the private sphere. Elson and Balakrishnan (2012) argue that this attitude has led to the neglect of policy to eliminate gender-based violence against women in the past. A lack of data collection should not, however, be a barrier towards transformative action. The continual existence of such an epidemic fundamentally undermines both women's equality and empowerment. As such, its inclusion in a post-2015 framework is strongly validated. It is less clear how mitigation of such violence can be achieved and how it should be framed to truly reflect gender equality progress.

No single intervention will eliminate violence against women. A combination of statistical indicators associated with infrastructure, legal, judicial enforcement, education, health and other service-related actions can significantly reduce it and its consequences. For such action to be taken, violence against women must be universally viewed as unacceptable. Its inclusion in a post-2015 framework would be a vital step in acknowledging this. The framing of such an issue should be reflected in the statistics outlined above, but accompanied by universally agreed upon data collection methods and research. Currently, the inclusion of gender-based violence in the post-2015 agenda does not directly challenge any of the criticisms associated with the de-contextualisation and oversimplification of gender equality as highlighted under MDG3. Given the lack of statistical reliability, if such diverse methods are not adopted, the complexities of gender-based violence are in danger of being subject to the same issues of obfuscating
progress as witnessed under MDG3. In responding to criticisms centred on MDG3, strong advocates propose a twin-track approach to gender equality in the post-2015 era.

**A Twin-Track Approach**

A twin-track approach to a post-2015 framework combines a stand-alone goal of gender equality with gender mainstreamed across all other objectives. Smee and Woodroffe (2013: 22), representing the Gender and Development Network, argue that this approach is the most likely to address the widespread structural inequalities, which persist for women and girls, and address issues associated with the oversimplification and de-contextualising nature of MDG3. A stand-alone goal on gender equality is undoubtedly necessary because gender equality is a goal in itself, deserving both focused political and financial commitment. Whilst its underlying goals and indicators must be reframed to better reflect contextual development progress, a stand-alone goal on gender provides clear lines of responsibility and a key advocacy lever for civil society to hold their national governments to account.

As an alternative to the twin-track approach, it has been proposed that gender equality should be grouped under a general inequalities theme across all new goals (Smee & Woodroffe 2013). Arguably, this strategy would undermine gender equality advocacy and would result in the scaling back of resources and political commitments (Spieldoch 2013). Although this may be true, gender mainstreaming has been widely criticised. Grown et al. (2005) argue that gender indicators that crosscut each goal are supposed to be everyone’s business, but they tend to become no one’s responsibility. This could lead to a lack of internal responsibility, an absence of accountability mechanisms, as well as resistance within organisational culture. Perhaps the main consequence of a mainstreaming agenda is that the guidelines for gender mainstreaming become an end in themselves rather than a process. This issue of policy evaporation could become a legitimate risk post-2015, as gender mainstreaming could become simply another requirement of funding. Advocates who propose the approach currently lack strong evidence to suggest this gender mainstreaming initiative would be successful and would challenge issues of de-contextualisation. Many seem more set on raising the profile of gender inequality through mainstreaming, without considering its potential negative consequences. Rather than gender becoming overtly mainstreamed across all goals, better collection and analysis of sex-disaggregated data underpinning each new target should be the priority. This should be accompanied by a stand-alone goal with reframed indicators, which focuses on the need to promote linkages with other global development objectives.
Conclusion

This essay argues that the framing of MDG3 to measure development progress obscures and oversimplifies the contextual nature of gender inequality. The indicators of progress adopted as part of MDG3: eliminating gender disparity in educational enrolment, increasing the share of women in non-agricultural wage employment and increasing the proportion of seats held by women in national parliaments, do have the potential to promote gender equality and empowerment. They are framed, however, through a set of reductionist variables that fundamentally obscure the development reality. The reframing of the post-2015 framework should build on the strengths of the current framework to find a balance between simplicity of communication for awareness purposes, while highlighting the complex and contextual nature of gender equality. Although important, the inclusion of gender-based violence as a post-2015 progress indicator does not directly challenge any of the criticisms currently underpinning MDG3. The lack of reliable statistical data suggests that the inclusion of this issue will inevitably be subject to similar criticisms of MDG3 to do with legitimacy and de-contextualisation. Additionally, the twin-track approach seems set on raising the profile of gender inequality through mainstreaming, yet does not challenge the issue of obfuscation. If direct efforts towards addressing such issues in policy formulation are not addressed, development progress will continue to be premised upon a de-contextualised and de-legitimatised definition of gender inequality.

MDG3 has been invaluable for the global gender equality movement. It has rallied unprecedented financial and political support, and introduced universally endorsed indicators that do have the potential to reduce inequality and enable empowerment. It is important, however, to remember the critical importance of this new framework for millions of women around the world who are subject to daily deprivation caused by gender inequality. Currently, the framing of MDG3 obscures the reality of such deprivation.

Bibliography


Domestic and Sexual Violence Related Asylum Claims in the Refugee Review Tribunal

Amanda Joyce Neilson

Domestic Violence-based Asylum Claims (DVAC) fall within a vexed area of Australian refugee law, which has been subject to significant academic and policy attention. Currently, an asylum bid by a victim of domestic violence for a Protection (Class XA) Visa turns on three elements: the membership of the claimant of a ‘particular social group’ (PSG), which is one of the five grounds listed as a prerequisite to refugee status in the 1951 Convention Relating to the Status of Refugees (‘the Refugees Convention’); a nexus between that PSG and the harm suffered by the claimant; and, an inability or unwillingness of the state in question to assist. The final requirement ensures that the harm meets the threshold set by the Refugees Convention, and is not a purely personal form of harm.

This essay reviews two decisions of the Refugee Review Tribunal, arguing that they demonstrate a reticence on the part of the Australian judiciary to provide a definitive legal ruling on the legitimacy of asylum claims for domestic violence, perpetrated by non-state actors. Using comparative legal analysis, the essay contends that while this approach is allied with the trends set by other Commonwealth jurisdictions, the Australian approach remains unique. It argues that the preference for guidelines and a policy of case-by-case assessment allows for a degree of flexibility in the assessment of claims in this area of refugee law to provide protection to those who may not have been able to seek it otherwise. The essay demonstrates, however, that this ambiguity can lead to practical difficulties for claimants in navigating such a broad scheme, particularly with regard to identifying their membership of a PSG and the giving of evidence in tribunal hearings.

Relevant Law and Issues to Consider

The relevant section of the Migration Act 1958 (Cth), s 36(2), provides that a Protection (Class XA) Visa may be granted to a ‘non-citizen in Australia to

---

whom the Minister is satisfied that Australia has protection obligations under the Refugees Convention as amended by the Refugees Protocol.\textsuperscript{2} Article 1A(2) of the Refugees Convention defines a refugee as someone who:

> owing to 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.\textsuperscript{3}

In order to be classified as a refugee, an applicant must be outside of his or her country, have a well-founded fear of persecution, fall into one of the five convention grounds and a nexus must exist between that ground and the persecution. Note that gender is not mentioned specifically in the convention.

**A The Definition of ‘Persecution’**

While the Refugees Convention does not provide a definition of persecution, it is widely acknowledged that the convention is ‘primarily aimed at the protection of individuals from the state of public forms of persecution, rather than intruding into the private realm’.\textsuperscript{4} Domestic violence is largely perpetrated by non-state actors for non-convention reasons, which raises the question as to whether victims of such violence are eligible to claim asylum under the Refugees Convention in the first place.

This question was answered by the High Court in *Minister for Immigration and Multicultural Affairs v Khawar*.\textsuperscript{5} In separate judgments, the majority agreed that where serious harm is inflicted by non-state actors for a non-convention reason, the nexus to the Refugees Convention will be met if the state in question withholds protection in a selective and discriminatory manner. Justice Kirby, following the decision of the English House of Lords in *Islam (AP) v Secretary of State for the Home Department Regina v Immigration Appeal Tribunal and Another Ex Parte Shah (AP) (Conjoined Appeals)*\textsuperscript{6} applied the formula, ‘Persecution = Serious Harm + The Failure of State Protection’.\textsuperscript{7} Chief Justice Gleeson held that persecution may be found where criminal conduct in a private context is condoned or tolerated by the state in situations where the state has a duty to provide protection from such conduct.\textsuperscript{8}

\textsuperscript{2} *Migration Act 1958* (Cth), s 36(2).
\textsuperscript{3} *Refugees Convention*, art. 1A(2).
\textsuperscript{5} (2002) 210 CLR 1 (‘Khawar’).
\textsuperscript{6} [1999] 2 AC 629 (‘Shah’).
\textsuperscript{7} *Khawar*, [118].
\textsuperscript{8} ibid., [30].
Post-\textit{Khawar}, the Migration Act was amended to insert s 91R, which purported to limit the definition of ‘persecution’ under the Refugees Convention. This amendment was made to ensure that the \textit{Khawar} decision did not widen the application of the Refugees Convention ‘beyond the bounds intended’. The section qualifies the term ‘persecution’, stating that Article 1A(2) does not apply in relation to persecution for a convention ground unless:

- the reason for the persecution is ‘the essential and significant reason’\textsuperscript{10}
- the persecution involves ‘significant harm to the person’,\textsuperscript{11} and
- the persecution ‘involves systematic and discriminatory conduct’.\textsuperscript{12}

The amendment also provided a list of examples of ‘serious harm’ to include a threat to a person’s life or liberty, significant physical harassment and/or ill-treatment, and economic hardship or denial of capacity to earn a livelihood that threatens a person’s capacity to subsist.\textsuperscript{13} The list is non-exhaustive, however, and thereby allows for a degree of judicial discretion, as demonstrated by the following case studies.

\section*{B Membership of a Particular Social Group}

In \textit{Applicant S v Minister for Immigration and Multicultural Affairs},\textsuperscript{14} the High Court defined a PSG as follows:

First, the group must be identifiable by a characteristic or attribute common to all members of the group. Secondly, the characteristic or attribute common to all members of the group cannot be the shared fear of persecution. Thirdly, the possession of that characteristic or attribute must distinguish the group from society at large.\textsuperscript{15}

In the case of family groups, whose claims rest on the fact that they themselves constitute a PSG, s 91S of the Migration Act states that a decision-maker must ‘disregard any fear of persecution … that any other or former member … of that family has ever experienced’, where that reason is not mentioned in Article 1A(2) of the Refugees Convention.\textsuperscript{16}

Whether a group is a PSG for the purposes of the Refugees Convention is dependent upon relevant information regarding legal, social, cultural and

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{9} Explanatory Memorandum, Migration Legislation Amendment Bill (No 6) 2001 (Cth), 18.
\item \textsuperscript{10} \textit{Migration Act 1958} (Cth), s 91R(1)(a).
\item \textsuperscript{11} ibid., s 91R(1)(b).
\item \textsuperscript{12} ibid., s 91R(1)(c).
\item \textsuperscript{13} \textit{Migration Act 1958} (Cth) s 91R(2).
\item \textsuperscript{14} (2004) 217 CLR 387.
\item \textsuperscript{15} ibid., [36] (Gleeson CJ, Gummow and Kirby JJ).
\item \textsuperscript{16} \textit{Migration Act 1958} (Cth) s 91S.
\end{itemize}
\end{footnotesize}
religious norms in their country of origin. Gender has previously been acknowledged by Australian courts as a PSG, however, the courts are swift to qualify these findings by stating that each case turns on its individual facts, and that analogy is not to be relied upon in the decision-making process.

C Nexus to the Refugee Convention

It is now well established in Australian jurisprudence that the nexus to the convention will be sufficient if a state is unable or unwilling to protect those suffering from persecution at the hands of non-state actors. Of course, this does not mean that every victim of domestic violence is now able to successfully obtain refugee status. As Bacon and Booth argue, ‘While violence against women may exist in most if not all cultures, its prevalence is certainly not uniform … and [it] certainly does not always receive the tacit acceptance of the state’. There is a significant evidentiary burden placed on an applicant, as ‘decision-makers are entitled to assume (unless the contrary is proved) that the state is capable within its jurisdiction of protecting an applicant’. In such situations, the courts must rely heavily on the socio-legal and political context and country information, alongside the particular claims of the applicant to make such determinations.

The Two Cases

Case 1: Domestic Violence in Papua New Guinea

The applicant applied for asylum in Australia to escape over 20 years of domestic violence and sexual abuse at the hands of her husband in Papua New Guinea. The applicant feared death or infection with HIV by her husband. She

17 Khawar, [26] (Gleeson CJ).
19 Khawar, [26] (Gleeson CJ).
20 ibid.
claimed that though she had sought assistance on a number of times previously, authorities in Papua New Guinea were consistently unwilling to protect her because she was a woman, and considered her husband’s property.25

The tribunal determined that the ‘violent beatings and sexual violence clearly satisfied the definition of “serious harm” amounting to persecution’26 under s 91R of the Migration Act. In their determination of the applicant’s status as a member of a PSG group, they referenced extensive country information to support the applicant’s claims. The tribunal determined that the country information ‘supports a broad interpretation of ‘women in PNG’ as a particular social group, ‘as the discriminatory and abusive treatment of women in PNG is pervasive’.27 They determined on the evidence that the applicant had been persecuted as a ‘woman in PNG’, a ‘woman from the Highlands’, a ‘married woman in PNG’, a ‘married woman in PNG for whom a bride price has been paid’ and as a ‘woman subject to domestic violence in PNG’.28 As the fear of harm at the hands of her husband was both continuing and not remote, the tribunal determined that the applicant met the convention nexus requirement.

Finally, the tribunal found that the applicant could not rely on state or police protection, noting the attitude of impunity towards domestic violence incidents in Papua New Guinea. This finding was based on country information, which indicated, ‘The police in PNG are unable to protect women and are complicit in the persecution of female victims of domestic violence’.29 The tribunal determined that the applicant should be granted refugee status on the basis of persecution for reasons of her membership of the particular social group ‘women in PNG’.30

Case 231: Interfamilial Rape in Fiji

The applicant, a Fijian-Indian woman, applied for asylum in Australia after being raped twice by her stepmother’s brother. She claimed that the authorities had taken no action to pursue her case, that her abuser had ‘contacts’ in the police force that prevented such action, and that she feared continuation of the sexual violence upon her return to Fiji as she had ‘nowhere else to go’ but back to her stepmother and abuser.32 The applicant did not submit as part of her claim that she was a member of any PSG.

25 ibid., [31].
26 ibid., [68].
27 ibid., [81].
28 ibid., [82].
29 ibid., [84].
30 ibid., [88].
32 ibid., [50].
The tribunal determined that ‘repeated sexual violence with intimidation and threats to kill the applicant should she report the crimes to the police’ constituted ‘significant ill-treatment of the person’, thus falling within the definition of serious harm, s 91R of the Migration Act. As the applicant’s stepmother’s brother perpetrated the rapes, the tribunal needed to seriously consider whether the harm was aimed at the applicant as an individual, and not for a convention reason. The tribunal determined that there was nothing to suggest that the rape was anything other than a ‘crime of opportunity perpetrated on the applicant as a weaker, more vulnerable person who happened to be home alone and unprotected’ at the time of the rapes.

The tribunal determined that the applicant was a member of the particular social group ‘unprotected single Indian-Fijian women with no family support’. Referencing extensive country information, the tribunal determined that the applicant’s gender, ethnicity and lack of familial support in Fiji were all significant reasons for her persecution. They acknowledged that it was difficult to make a definitive finding on whether police protection would be withheld from the applicant. They determined, however, that from the evidence presented, ‘the current authorities of Fiji will not afford the applicant effective state protection from the threat of sexually violent crimes’. The applicant was therefore granted asylum on the ground of her fear of serious physical harm for reason of her membership of the particular social group: ‘unprotected single Indian-Fijian women with no family support’.

**Comparative Analysis of the Australian Approach**

**A Persecution, the Convention Nexus and the Public/Private Dichotomy in Refugee Law**

The Australian courts have demonstrated a willingness to take a holistic approach to the reconciliation of domestic violence-related asylum claims with the requirements of the Refugees Convention. Domestic violence is generally perpetrated in the private sphere, which on face value creates difficulty in reconciling the private nature of the harm with the traditionally public nature of persecution imagined under the Refugees Convention. The *Khawar* decision emulates the broader jurisprudential trend throughout Commonwealth countries.

---

33 ibid., [74].
34 ibid., [89].
that persecution at the hands of non-state actors will fall within the convention requirements ‘in situations where the State is unable or unwilling to afford effective protection against such harm’.36 This acknowledges that persecution is not exclusive to the public sphere, and that private harm by non-state actors can lead to serious abuse of the human rights of victims of domestic violence.

The landmark ruling on this issue was made by the House of Lords in Shah.37 In this case, two married Pakistani women sought asylum in the United Kingdom for fear that, if they returned to Pakistan, they would be beaten by their husbands and potentially prosecuted for adultery. The penalties for adultery include flogging or being stoned to death. Lords Hoffman and Steyn held that harm caused by a non-state actor could constitute persecution where there was an inability or unwillingness on the part of the State to provide effective protection, for reasons of membership of a gender-defined particular social group.38 Steyn stated, ‘Given the central feature of state-tolerated and state-sanctioned gender discrimination, the argument that the appellants fear persecution not because of membership of a particular social group but because of the hostility of their husbands is unrealistic’.39

The Shah precedent was refined by a decision of the New Zealand Refugee Status Appeals Authority (RSAA) in 2000. In Refugee Appeal No. 71427/99, Haines QC found that the applicant, an Iranian woman seeking asylum for fear of further violence from her abusive husband, was trapped in a system of ‘state-sanctioned inequality … in the context of the Islamic theocratic regime’,40 and that this constituted persecution within the confines of the Refugees Convention. Haines QC ‘unpacked and made explicit many of the underlying assumptions of Shah, creating a framework both practically and theoretically coherent for analysing persecution and the relationship between non-State actor agency and state responsibility’.41 The decision continued the trend of applying a ‘holistic understanding of refugee status and protection’42 within the context of domestic violence-related asylum claims.

In Khawar, the High Court acknowledged both Shah and Refugee Appeal No. 71427/99 in its decision to grant asylum to the appellant. Of particular importance is the judgment of Gleeson CJ, who stated that persecution can

37 [1999] 2 AC 269 (‘Shah’).
38 ibid., 558 (Lord Steyn); 564–65 (Lord Hoffman).
39 ibid., 558 (Lord Steyn).
40 Refugee Status Appeals Authority (RSAA), 16 August 2000, [7].
42 ibid., 397.
be based on the ‘combined effect’ of multiple agents. He reasoned that in the context of state-sanctioned domestic violence — which *prima facie* may be perpetrated for a non-convention reason — the nexus requirement will be satisfied if at least one agent of persecution (state or individual) persecutes the victim for a convention reason. He stated:

Where persecution consists of two elements, the criminal conduct of private citizens, and the toleration or condonation of such conduct by the state or agents of the state, resulting in the withholding of protection which the victims are entitled to expect, then the requirement that the persecution be by reason of one of the Convention grounds may be satisfied by the motivation of either the criminals or the state.

The willingness of the Australian Refugee Review Tribunal to take full advantage of this broad mandate is evidenced by the case studies presented above. In Case 1 the systemic nature of both the abuse and refusal of PNG authorities to provide protection to the applicant meant that her claim can easily fit into the framework above. Case 2 proved more problematic for the tribunal in its decision as to whether the applicant met the persecution and nexus requirements for asylum. In Case 2, the tribunal granted asylum to the applicant, even in the face of uncertainty as to whether the Fijian police force would in fact refuse to take action against her abuser. They considered not only her immediate reasons for seeking asylum, but also her personal situation and the context in which she was seen in Fijian society as a whole. The PSG to which she was ascribed — an ‘unprotected single Indian-Fijian women with no family support’ — intimates that the tribunal sought to combine gender, race and personal circumstances in order to create a situation in which each of the convention requirements were met.

The precedents set in the above cases have also, notably, begun to permeate the bureaucratic processes before asylum claims reach the tribunal stage. The recent revision of the Department of Immigration and Border Control’s guidelines for assessing gender-related asylum claims now acknowledges that ‘Sexual violence and harm perpetrated in the “private” sphere or by non-state agents can … amount to persecution’, and that ‘state protection may be withheld for a Convention reason … [which can] amount to persecution under the Refugees Convention’.

43 *Khawar*, [27] [Gleeson CJ].
44 ibid., [31] [Gleeson CJ].
B The Particular Social Group Convention Ground

The PSG convention ground is ‘regarded as the most amorphous of all the Convention grounds’, and has proven problematic in both domestic and international jurisprudence. Australia has favoured a broad approach when interpreting the ground, following decisions from other Commonwealth jurisdictions including Canada and New Zealand. Each of these jurisdictions’ interpretations of the convention ground have reflected the two ‘dominant approaches’ espoused by the United Nations Commissioner for Refugees (UNHCR) in its guidelines for the interpretation of PSG.

The UNHCR guidelines for the interpretation of PSG recommend that the following definition should be adopted:

A particular social group is a group of persons who share a common characteristic other than their risk of being persecuted, or who are perceived as a group by society. The characteristic will often be one which is innate, unchangeable, or which is otherwise fundamental to identity, conscience or the exercise of one’s human rights.

The two dominant approaches are summarised by Kneebone as the ‘protected characteristics’ approach and the ‘social perception’ approach, which ‘recognise respectively both the internal unifying characteristics of the group and the external perception or visibility of the group’.

Both Canada and New Zealand have applied a broad interpretation of a PSG to include women fleeing persecution and women of Iran, for example. In Canada (Attorney General) v Ward the Supreme Court of Canada outlined three categories of PSG. These were (1) ‘those defined by an innate or unchangeable characteristic, for example individuals fearing persecution by reason of gender’; (2) ‘those whose members voluntarily associate for reasons so fundamental to their human dignity that they should not be forced to forsake the association’; and, (3) ‘those associated by a former voluntary status, unalterable due to its historical permanence’.

---

49 Kneebone, ‘Women Within the Refugee Construct’, 27.
50 Mayers and Marcel v MEI (1992), 97 DLR (4th) 729 (FCAD).
52 ibid., 739.
Similarly, in *Refugee Appeal no 71427/99*, the NZ RSAA held that ‘women of Iran’ were a PSG on the basis that the ‘overarching characteristic of those fundamentally disenfranchised and marginalised by the state is the fact that they are women’. Mullally argues that by broadly interpreting the convention to include ‘women of Iran’, ‘[t]he case goes beyond the more limiting, and in some cases tautological, claims that the social group must be limited further, for example to women who are at risk of spousal or intimate partner abuse’. She states that in assessing the reasons for persecution in this case, ‘[g]ender is recognised as the relevant variable’.

In *Khawar*, the majority agreed that ‘women in Pakistan’ constituted a particular social group; yet each judge favoured a slightly different approach. Gleeson CJ made the broadest interpretation, arguing that there was ‘nothing inherently implausible in the suggestion that women in a particular country may constitute a persecuted group’. Kirby J stated that ‘[t]here may be a particularly vulnerable group of married women in Pakistan, in dispute with their husbands … unable to call on male support’, while McHugh and Gummow JJ suggested that at its narrowest, the group might be categorised as ‘married women in a household that did not include a male blood relation to whom the woman might look for protection against violence by the members of the household’.

Crucially, Gleeson CJ emphasised that each decision must be decided on its facts, and that the decision-makers must take into account the particular context within which the asylum claim is made. He states:

> An Australian court or tribunal would need to be well-informed about the relevant fact and circumstances, including cultural conditions, before reaching a conclusion that what occurs in another country amounts to persecution by reasons of the attitudes of the authorities to the behaviour of private individuals.

The High Court’s reluctance to narrow the definition of PSG and emphasis on case-by-case assessment practice has allowed lower courts and tribunals to make nuanced assessments on the basis of the facts of each case, while managing the legislative requirements under the Migration Act. This is reflected in the two case studies above. In Case 1, even though narrower potential PSGs were presented

---

54 *ibid.*
55 *ibid.*, [108].
57 *ibid.*
58 *Khawar*, [33] (Gleeson CJ).
59 *ibid.*, [129] (Kirby J).
60 *ibid.*, [81] (McHugh and Gummow JJ).
61 *ibid.*, [26] (Gleeson CJ).
to the tribunal, it granted asylum based on the applicant’s membership of the particular social group ‘women in PNG’. This was largely based on country information citing the pervasive gender discrimination and tacit acceptance of domestic violence by PNG authorities.

In Case 2, it was necessary to tailor the concept of PSG to the facts of the case, as the evidence required to claim a broader PSG, as in Case 1, simply did not exist. The tribunal relied on in-country reports from Amnesty International and US State Departments, alongside local news reports covering incidents of domestic violence, incest and racial discrimination of Indian-Fijians to piece together the extremely narrow PSG of ‘unprotected single Indian Fijian women with no family support’. This was despite the applicant herself failing to submit a claim to a PSG in the first instance, demonstrating a willingness on the part of the Refugee Review Tribunal to assist victims of domestic violence with their asylum claims to the limit of the legislative requirements.

C Practical Challenges of the Australian Approach

Australia’s broad approach to assessing asylum claims made by victims of domestic violence, while simultaneously widening the scope for such claims has created a number of practical difficulties for claimants. The case-by-case assessment policy requires extensive research by the decision-maker or tribunal into both the personal situation of the applicant and the context in which they are making their claim. The Australian Law Reform Commission (ALRC) states that such a rigorous process ‘may be difficult for those who face language barriers, lack legal representation, or lack access to current country information’.62

Regarding membership of a PSG, the ALRC emphasises that applicants are forced to tread a line between claims that are too broad, risking ‘a finding that the harm feared is not motivated by their membership of a particular social group’, or too narrow —risking a determination ‘that the group is impermissible as it is defined by the harm feared.’63 In Case 2, for example, the applicant’s lack of access to legal representation meant that she was not able to submit her membership of a PSG to the tribunal at all; it had to be imputed by the tribunals on the facts of the case.64

For victims of domestic violence, recalling in detail the incidents that led to their claim may lead to an exacerbation of the trauma stemming from the original harm. In Case 2, to elucidate the facts of the claim, the tribunal needed to ask questions, including an explanation of the applicant’s claims of rape,65

---

63 ibid.
64 1002732 [2012] RRTA 365 (7 February 2012) [41].
65 ibid.
whether the applicant was raped by more than one person,\textsuperscript{66} an account of the first rape, and her activities immediately following both rapes.\textsuperscript{67} These questions proved so traumatic for the applicant that the tribunal adjourned the proceedings on four separate occasions to allow the applicant time to recover. The tribunal was presented with a professional psychological assessment stating that the applicant was suffering from post-traumatic stress disorder as a result of the abuse suffered,\textsuperscript{68} the recounting of which only caused further distress. Yet the tribunal apparently needed to ask these probing questions to assess the applicant’s case, given the case-by-case assessment mandate proffered by Gleeson CJ in \textit{Khawar}. This ‘necessary evil’ is a demonstration of the practical implications of decisions at the highest level — a flaw in a system designed to protect victims of domestic violence, exposing them to further trauma in order to have their claims assessed.

## Conclusion

Instances of domestic and sexual violence by non-state actors where the state is unable or unwilling to provide effective protection to the victims are well recognised as bases for claims for asylum in Australia. The Australian judiciary is ambiguous as to any definitive ruling on the legitimacy of such claims, preferring the setting of guidelines, and a case-by-case assessment policy. This ambiguity has allowed for a degree of flexibility for decision-makers and lower courts and tribunals to grant asylum to those who may not have been able to seek it otherwise.

The key criticism with this approach is its practical effects on those claiming asylum for domestic and sexual violence, as the case-by-case assessment mandate means that the victims must once again revisit their reasons for claiming asylum in order to seek protection under Australian refugee law.

## Bibliography

### Articles/Books/Reports


\textsuperscript{66} ibid.
\textsuperscript{67} ibid., [46].
\textsuperscript{68} ibid., [60].


**Cases**

Applicant A v Minister for Immigration and Ethnic Affairs (1997) 190 CLR 22.
Islam (AP) v Secretary of State for the Home Department Regina v Immigration Appeal Tribunal and Another Ex Parte Shah (AP) (Conjoined Appeals) [1999] 2 AC 629.

Mayers and Marcel v MEI (1992), 97 DLR (4th) 729 (FCAD).
SZAIX v Minister for Immigration and Multicultural and Immigrant Affairs [2004] FCA 104.

Legislation
Migration Act 1958 (Cth).

Treaties

Other
Department of Immigration and Border Security, ‘PAM3 Guidelines: Gender-Related Asylum Claims’.
Explanatory Memorandum, Migration Legislation Amendment Bill (No. 6) 2001 (Cth).

Susan Yuan Wen Nheu

Abstract

L’Enfant et les sortilèges is a lyric fantasy in one act by French composer Maurice Ravel and French writer and librettist Sidonie-Gabrielle Colette. This work, one of Ravel’s two contributions to the lyric stage, was first performed at Monte Carlo on 21 March 1925 and, one year later, it premiered at the Opéra-Comique in Paris. Sparsely documented communication and personal contact in the nine years between the work’s conception and premiere seems to indicate that there was little collaboration on the work. Through a study of primary and secondary sources and an examination of the strong text–music relationship, however, I argue that L’Enfant was in fact the result of a genuine collaboration between Colette and Ravel. My argument is supported by a study of the genesis of the opera, Ravel’s manipulation of the rhymes and rhythms of Colette’s libretto, the relationship the work has with fairytale narratives and storytelling, and Ravel’s unique compositional technique.

The text and music in L’Enfant et les sortilèges is the result of one of the greatest collaborations of the 1920s. This work is a lyric fantasy in one act and is one of Maurice Ravel’s two contributions to the lyric stage. It was first performed at Monte Carlo on 21 March 1925 under the baton of Victor de Sabata; one year later, it premiered at the Opéra-Comique in Paris.¹ L’Enfant is the story of a rude and disrespectful child who disobeys his mother’s wishes and goes on a destructive rampage in his room. The cast includes fairytale characters, crockery, furniture, a homework book, plants and small animals, which come to life to reprimand the child who has been tormenting and hurting them. Through his spiritual journey, the child experiences love, friendship and forgiveness as he develops his moral compass and awareness of the consequences of his behaviour. At a time when successful collaborations between composers and

writers of stature were surprisingly rare.\(^2\) L'Enfant was unique in pairing the most acclaimed living composer in France, Ravel,\(^1\) with Sidonie-Gabrielle Colette,\(^4\) a French writer who was equally recognised as one of the finest of her generation', to create an opera-ballet set to a new libretto.\(^5\) During the nine years between the work’s conception and premiere (1916–25), the dearth of documented communication between the composer and librettist gives rise to the idea that there was little or no truly ‘collaborative’ process in the preparation of L'Enfant.\(^6\) Through exploring the genesis of the opera, Ravel's manipulation of the rhymes and rhythms of Colette's libretto, the relationship the work has with fairytale narratives and Ravel's compositional technique, I argue that, in contrast, despite having worked mostly independently, Ravel and Colette were genuine collaborators.

Understanding the genesis of L'Enfant illuminates the events and circumstances of the composer's life, which interrupted and delayed L'Enfant's composition. The genesis of the opera can be traced back to World War I.\(^7\) In 1916, Jacques Rouché, director of the Opéra Garnier, approached Colette with a commission for a libretto of a 'fairy-ballet'.\(^8\) Following Colette's agreement to write the libretto, and having been turned down by composer Paul Dukas, Rouché approached Ravel with the commission. Colette sent Ravel a script in late 1916 while he was stationed at Verdun, but, as a result of the outbreak of war, it did not reach him.\(^9\) Despite this first delay, Ravel eventually received the draft of

---

3. French composer Maurice Ravel (1875–1937) was one of the most original and sophisticated musicians of the early 20th century. He was a master of orchestral and pianistic tone colour. Although he is often described as an ‘impressionist’ and paired with his contemporary, Claude Debussy, Ravel does not fit neatly into any musical style category — ‘much of his music is too clearly defined in form and tonality and classically balanced in phrase structure to be considered impressionist’. Roger Kamien, Music An Appreciation, New York: McGraw-Hill, 2004. Barbara L. Kelly, ‘Ravel, (Joseph) Maurice’, Grove Music Online, Oxford University Press, accessed 29 October 2013, http://www.oxfordmusiconline.com/subscriber/article/grove/music/O005396.
4. French writer Sidonie-Gabrielle Colette (1873–1954) became well-known in Parisian musical circles after her marriage to the music critic known as Willy (Henri Gauthier-Villars). According to The New Grove Dictionary of Opera, ‘her “Claudine” novels — which Willy published under his own name — created a vogue, one result of which was the operetta Claudine by Rodolphe Berger.’ Colette also worked as a music critic, and on several occasions she was sent to the same concert as Debussy to produce a different viewpoint. Richard L. Smith, ‘Colette’, The New Grove Dictionary of Opera, Grove Music Online, Oxford University Press, accessed October 29, 2013, http://www.oxfordmusiconline.com/subscriber/article/grove/music/O005396.
the libretto in 1918, however, there is no record of him working with it until 1920. According to the first known direct communication between Ravel and Colette, another delay was Ravel’s poor health. ‘My only excuse,’ Ravel wrote, ‘is the state of my health: for a long time, I feared that I would never be able to do anything.’

Furthermore, following the death of Ravel’s mother in January 1917, he was in no state to undertake a major project. This period of grief marked a turning point in Ravel’s career, according to Arbie Orenstein:

its immediate effect was some three years of virtual silence with regard to composition and from this point until his swan song … only about one composition a year would be completed.

It is, therefore, evident that Ravel’s life circumstances delayed L’Enfant’s composition.

Primary sources reveal that a collaborative approach towards L’Enfant was in existence from the very beginning and the artists developed a shared understanding on the project. The correspondence between Ravel and Colette is evidence of an exchange of ideas. For example, when Ravel asked, ‘What would you think of the cup and teapot, in old Wedgwood — black — singing a ragtime?’, Colette responded, ‘Why certainly a ragtime! Why of course Negroes in Wedgwood!’ It is clear that Ravel valued Colette’s input and ideas when he states, ‘I would be grateful to have your opinion on these two points …’. One week following Ravel’s letter, Colette wrote in her return letter that ‘the squirrel will say everything you wish’. She also seeks for Ravel’s artistic judgment, ‘Does the “cat” duo, exclusively meowed, please you?’. A spirit of camaraderie, collaborative enthusiasm and a shared sense of artistic sympathy between Ravel and Colette is unmistakeable in their correspondence.

The word and essence of collaboration carries on throughout their correspondence. The choice of language employed by both Ravel and Colette conveys a sense of artistic partnership: ‘our opera’; refers to the existence of a collaboration: ‘to see if she still wants to have me as a collaborator’; and

10 Kilpatrick, ‘Enchantments and Illusions’, p. 36.
11 Orenstein, Ravel, p. 78.
12 Kilpatrick, ‘Enchantments and illusions’, p. 36.
13 Arbie Orenstein, A Ravel Reader: Correspondence, Articles, Interviews, Mineola, New York: Dover Publications, 2003, p. 188. Refer to appendices A and B.
14 Kilpatrick, ‘Enchantments and illusions’, p. 34.
15 ibid.
16 Orenstein, Ravel: Man and Musician, p. 75.
17 Orenstein, A Ravel Reader, p. 188.
18 ibid.
is punctuated with words of encouragement and support for each other’s roles: ‘Go for it!’\textsuperscript{20} It is clear that the artists intended \textit{L’Enfant} to be a shared project and that Ravel and Colette clearly supported and respected each other’s roles as collaborators. Despite the events and circumstances that delayed the composition, the primary sources are explicit in revealing that Ravel and Colette agreed to work on \textit{L’Enfant} as a collaborative project from the very beginning.

Ravel and Colette shared an artistic affinity for the interplay of music and language. As a composer, Ravel gave much consideration to the words in \textit{L’Enfant} through his manipulation of the rhythms and resonances of the language; he heightened the dramatic and musical effects of Colette’s words. Ravel’s input in linguistic effect can be seen in the textual coherence between the contrasting episodes in the opera. From the beginning to the last scene, the same poetic gesture is used to generate the same distinctive effects creating subtle onomatopoeia. As observed by Marie-Pierre Lussus, ‘many characters of the individual “enchantments” (“sortilèges”) are all shaped by the sounds of their words as much as by their meanings’.\textsuperscript{21} An example is the grouping of the vowel and consonant sounds of the sortilèges into two categories to emphasise the differences between the ‘hard’ and ‘soft’ sounds. The princess, the cats, the dragonfly and the squirrel are grouped in the ‘soft’ sound category while the other sortilèges — the child’s opening monologue, the clock, the arithmetic scene, the fire and the teapot — have ‘hard’ vowel and consonant sounds. For example, the ‘hard’ sounds are represented by the child singing ‘\textit{J’ai pas envie de faire mapage}’, and by the teapot in the opening scene with the line ‘Black and costaud, black and chic …’,\textsuperscript{22} while the ‘soft’ sounds, such as b/d/g, the fricative z [rose], ε [aime and cherche], and the slow θ [jeu], are used in the libretto sung by the dragonfly. As seen in Appendix D, the contrast between these two linguistic affects can be seen from the libretto of the fire and the dragonfly:

The fire

\textit{Je rechauffe les bons; Je brûle les méchants; Petit barbare, barbare imprudent; Tu as insulté tous les dieux bienveillant; Quitendent entre le malheur et toi; Le fragile barrière! …; Gare au feu dansant!; Gare, gare, gare, gare à toi!}\textsuperscript{23}

\textsuperscript{20} Orenstein, \textit{A Ravel Reader}, p. 188.
\textsuperscript{21} Kilpatrick, ‘Enchantments and Illusions’, p. 41.
\textsuperscript{22} ibid., p. 37.
The dragonfly

Où es tu?; Je te cherche; Le filet; Il t’a prise; O toi chère; Longue et frère;
Tes turquoises; Tes topazes; L’air qui t’aime; Les regretted; Moins que
moi …24

The subtle onomatopoeia achieved through the emphasis and juxtaposition of
‘hard’ and ‘soft’ vowel and consonant sounds gives the sortilèges an additional
depth of character.25 Clearly, Ravel gave much consideration to the words in
L’Enfant and he manipulated the qualities of Colette’s text to maximise both its
aural and dramatic effects in the opera.26

Furthermore, given Colette’s training as a pianist in her youth and her work in
musical criticism for example, of Gil Blas,27 her musical background was certainly
a factor contributing to the development of a smooth prosaic style, which played
with the expression of rhythm and assonances.28 Considering Colette’s musical
background, L’Enfant’s textual coherence matches well with a poet who plays
with words and sounds to bring out the dramatic effects of Ravel’s music. As we
see in the authorship of the lines and phrases, L’Enfant exudes a strong sense of
artistic affinity between composer and librettist through its musical and poetic
interplay. As observed by Raymond Balliman, the text and music of L’Enfant
were so intertwined that ‘it is impossible to separate the collaborators’.29 Ravel
and Colette shared a truly interdisciplinary sense of what an operatic libretto
might be in L’Enfant, where the words were shaped by musical sense and music
by textual sonorities.30

The fantasy and fairytale theme of L’Enfant suggests that, through the
expressive use of language, Ravel and Colette were deliberately aligning music
and text with the tradition of storytelling, a genre defined by the sounds of
the spoken word.31 Many of the hallmarks of the fairytale narrative are present
in L’Enfant: ‘a timeless setting; unnamed characters (the child, the fire, the
princess), reciprocal events and actions (the destructive child is himself injured
and becomes “good”), a happy ending that suggests a clear moral, and a
narrative shaped by “magical” events’.32 L’Enfant’s connection to storytelling
further supports the strong text–music relationship and collaborative nature
of the work. Ravel and Colette were clearly aware of the subtleties of sounds

24 ibid.
25 ibid., p. 10.
26 ibid., pp. 10–11.
29 ibid., p. 42.
30 ibid., p. 46.
31 ibid., p. 1.
32 ibid., p. 9.
and language in their respective métier and, through their collaborative efforts, they acted as storytellers to achieve the regular and natural rhythmic patterns, illustrative rhymes and assonances, and expressive inflections characteristic of fairytale narratives. An example can be seen in the ‘Cats’ Duet’, for which Ravel contributes to the phonetic representation of the cats mating scene ‘maou’ and also in the croaking frog’s line ‘Kékékékécékéça’ in the garden scene. Through the expressive use of language combined with musical effect, Ravel and Colette worked together to achieve a sense of storytelling in L’Enfant.

Collaboration on L’Enfant is also demonstrated by the relationship between Ravel’s composition technique and its function in assisting with the communication of Colette’s text. Throughout the opera, as seen in the development of the main protagonist (the child), Ravel employs begin-end transformation for the purpose of creating isomorphic musical transformations within the text. For example, in the beginning state, the child is naughty and he goes against Maman’s (‘Mother’s’) wishes through his destructive rampage against the animals and objects in the house. As seen in Appendix F, in this scene, a series of transpositions convey the opposition between the mother and child. The progressively higher transpositions of Mother’s cadence (MC) reflect the child’s willful silence in response to his mother’s queries and also her mounting displeasure. The scene concludes with a final statement of Mother’s cadence in the original pitch, giving the last word to Maman. There is an isomorphic representation in the music where the untransposed MC conveys the mother’s unshakeable moral authority. After the child’s rampage in the room, the animals and objects come to life and reveal to the child the consequences of his behaviour. In the last stage, the child embraces Maman and morality. The child’s transformation is complete as he sings Maman accompanied by MC. The MC as a perfect cadence communicates the strength and finality of Maman’s will, and the child’s submission to it. Here, the accompaniment by the MC

33 ibid., p. 10.
34 ibid., p. 12.
36 ‘Transposition’ refers to ‘the notation or performance of music at a pitch different from that in which it was originally conceived or notated, by raising or lowering all the notes in it by a given interval.’ Julian Rushton, ‘Transposition’, Grove Music Online, Oxford Music Online, Oxford University Press, accessed 29 October 2013, http://www.oxfordmusiconline.com/subscriber/article/grove/music/28274.
37 ‘The Mother’s cadence’ (MC) refers to the ambiguous cadential figure during the last restatement of the prelude that immediately precedes the entrance of Maman. See Appendix E for more detailed analysis. ‘Cadence’ refers to ‘the conclusion to a phrase, movement or piece based on a recognizable melodic formula, harmonic progression or dissonance resolution; the formula on which such a conclusion is based’. William S. Rockstro, et al., ‘Cadence’, Grove Music Online. Oxford Music Online. Oxford University Press, accessed 29 October 2013, http://www.oxfordmusiconline.com/subscriber/article/grove/music/04523.
38 ibid.
39 ibid., p. 41. See Appendix I to see the child’s transformation.
40 ibid.
41 ibid.
isomorphically represents the child’s acceptance of Maman’s moral authority. Ravel’s composition method creates the musical and dramatic effects to convey the meaning and emotion in the text. For example, Maman’s cadence serves as the agent for the child’s transformation from bad to good. Although Maman is not literally present on stage, her cadence is restated in each opening and closing scene with musical changes in key, orchestration and chord to create different dramatic musical effects that relate to the developing plot. The isomorphic musical transformations in *L’Enfant* are important in communicating Colette’s writing. Conversely, without Colette’s words, the isomorphic associations and meaning would be absent from the music. Ravel’s unique composition technique clearly demonstrates the close relationship that music has with the text in representing the meaning and structure of the work.

In conclusion, through exploring the genesis of *L’Enfant*, studying Ravel’s manipulation of the rhymes and rhythms of Colette’s libretto, the relationship *L’Enfant* has with fairytale narratives, and Ravel’s unique composition technique, there is much evidence to support the argument that, despite the lack of recorded correspondence between the composer and librettist during 1916–25, *L’Enfant* was nonetheless the result of a genuine collaboration between Ravel and Colette. As Ravel said himself, the nature of collaboration in *L’Enfant* lies between the ‘keen’ artistic sympathies shared by the two individuals.42 While they worked in different métiers, both Ravel and Colette were storytellers, and it is in this act of storytelling that Ravel and Colette best expressed their keen artistic sympathy.

**Appendix A — The first known direct communication between the collaborators;43**

Ravel to Colette on 27 February 1919.

February 27, 1919

Dear Madame,

While you were expressing your regrets to Rouché about my silence, I was thinking, amid all the snow, of asking you if you still wanted such an unreliable collaborator.

My only excuse is the state of my health: for a long time, I feared that I would never be able to do anything. It appears that I’m improving: the desire to work seems to be returning. Here, it’s not possible; but as soon as I return, at the beginning of April, I’m planning to resume work, and will begin with our opera.

---

42 Orenstein, *A Ravel Reader*, p. 188.
43 Kilpatrick, ‘Enchantments and Illusions’, p. 36.
In truth, I am already working on it: I’m taking notes — without writing any —;

I’m even thinking of some modifications … Don’t be afraid: they’re not cuts — on the contrary. For example: couldn’t the squirrel’s dialogue be extended? Imagine everything that a squirrel could say about the forest, and how that could be interpreted in music!

Another thing: what would you think of the cup and teapot, in old Wedgwood — black — singing a ragtime? I confess that the idea of having a ragtime sung by two Negroes at our National Academy of Music fills me with great joy. Note that the form — a single couplet, with refrain — would be perfectly suited to the gestures in this scene: complaints, recriminations, furor, pursuit. Perhaps you will object that you don’t usually write American Negro slang. I, who don’t know a word of English, will do just like you: I’ll work it out.

I would be grateful to have your opinion on these two points, and, dear Madame, believe in the keen artistic sympathy of yours truly,

Maurice Ravel

Source: Arbie Orenstein, _A Ravel Reader: Correspondence, Articles, Interviews_, Mineola, New York: Dover Publications, 2003, p. 188.

**Appendix B — Colette’s response to Ravel on 5 March 1919**

March 5, 1919

69 Boulevard Suchet

Paris XVI

Dear Sir,

Why certainly a ragtime! Why of course Negroes in Wedgwood! What a terrific gust from the music hall to stir up the dust of the Opéra! Go to it! I am glad to know that you are still thinking of ‘Divertissement for my Daughter’. I despaired of you, and was told that you were sick. Do you know that orchestras in cinema houses are playing your charming _Mother Goose_ suite, while they show American westerns? If I were a composer and Ravel, I think I would derive much pleasure from learning that.

And the squirrel will say everything you wish. Does the ‘cat’ duo, exclusively meowed, please you? We’ll get acrobats. Isn’t the Arithmetic business a polka?
A Genuine Collaboration in Text and Music

I wish you good health, and shake your hand impatiently.

Colette de Jouvenel


Appendix C — The *sortilèges* are grouped according to their vowel and consonant sounds into two categories:

<table>
<thead>
<tr>
<th><em>Sortilèges</em> characters with ‘hard’ vowel and consonant sounds</th>
<th><em>Sortilèges</em> characters with ‘soft’ vowel and consonant sounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>The child’s opening monologue, the clock, the arithmetic scene, the <em>rire</em> and the teapot.</td>
<td>The princess, the cats, the dragonfly and the squirrel.</td>
</tr>
</tbody>
</table>

Examples:
Make more use of the hard ‘a’ and incisive ‘e’ vowels, together with the consonant sounds f, t and z.
1) The child sings ‘*J’ai pas envie de faire mapage*’
2) The teapot’s scene opens with the line ‘Black and costaud, black and chic …’

Examples:
Make more use of ‘soft’ sounds: the consonants b/d/g, the fricative *z* [rose], the light i [libre] and *ɛ* [aime and cherche], and the slow 0 [jue] œ [coeur] and o [rose] vowels.
1) The squirrel sings ‘*mes beaux yeux*’
2) The dragonfly’s line ‘*Je te cherche*’


Appendix D — The contrast between these two linguistic affects can be seen in comparison of the fire’s libretto with the dragonfly’s.

<table>
<thead>
<tr>
<th>[The Fire]</th>
<th>[The Dragonfly]</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Je rechauffe les bons</em></td>
<td>Où es tu?</td>
</tr>
<tr>
<td><em>Je brûle les méchants</em></td>
<td><em>Je te cherche</em></td>
</tr>
<tr>
<td><em>Petit barbare, barbare imprudent</em></td>
<td><em>Le filet</em></td>
</tr>
<tr>
<td><em>Tu as insulté tous les dieux bienveillant</em></td>
<td><em>T’a prise</em></td>
</tr>
<tr>
<td><em>Quindent entre le malheur et toi!</em></td>
<td><em>O toi chère</em></td>
</tr>
<tr>
<td><em>Le fragile barrière!</em>…</td>
<td><em>Longue et frêle</em></td>
</tr>
<tr>
<td><em>Gare au feu dansant!</em></td>
<td><em>Tes turquoises</em></td>
</tr>
<tr>
<td><em>Gare, gare, gare à toi!</em></td>
<td><em>Tes topazes</em></td>
</tr>
<tr>
<td><em>L’air qui t’aime</em></td>
<td><em>L’air qui t’aime</em></td>
</tr>
<tr>
<td><em>Les regrette</em></td>
<td><em>Les regrette</em></td>
</tr>
<tr>
<td><em>Moins que moi…</em></td>
<td><em>Moins que moi…</em></td>
</tr>
</tbody>
</table>

Appendices E — I: A musical analysis of the child’s development through three stages.

Appendix E: Pentatonic variations and ‘moral choice’

Ex. 2 *L’enfant et les sortilèges*

a) Pentatonic variations and ‘moral choice’

<table>
<thead>
<tr>
<th>Statement</th>
<th>1a</th>
<th>1b</th>
<th>2a</th>
<th>2b</th>
<th>3a</th>
<th>3b</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 bars</td>
<td>7 bars</td>
<td>2 voices (2 ob.)</td>
<td>same</td>
<td>3 voices (add cb)</td>
<td>same</td>
<td>4 voices (add child)</td>
</tr>
<tr>
<td>G pentatonic</td>
<td>(G A B D E (F#))</td>
<td>add F♭</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Pentatonic variations and ‘moral choice’


As suggested by the innocent sound of two oboes, the openness of the parallel 5th and 4th, and the tonal neutrality of the pentatonic pitch collection (G–A–B–D–E–[F#]), it may be said that the music conjures up a world of moral possibilities and choices to the young child. Three restatements of the prelude are given and, during the last restatement, there is an interruption by an ambiguous cadential figure. This immediately precedes the entrance of Maman and it can be known as MC (mother’s cadence).
Appendix F—Maman’s moral authority and the child’s disobedience

Figure 2: Maman’s moral authority and the child’s disobedience

Appendix G — The penultimate scene prior to the child’s actual transformation stage 1

Ex. 2 (continued)

c) Maman’s cadence, tonal context, child’s transformation – stage 1

Figure 3: The penultimate scene prior to the child’s actual transformation stage 1


In this scene, the animals bear witness to the child’s goodness, but they are powerless to dress the child’s wound. They realise that only Maman can heal the child, and they call for her: ‘Maman! Maman! Maman!’ Although not literally present on the stage, Maman makes a dramatic return via the MC. Some musical
changes in the MC include: a different orchestration and the cadence has a
definite tonal context — functioning as ii–V in G major and resolving to the
tonic, coinciding with ‘Il est bon.’ ‘This suggests that through this change in the
tonal context, the latent moral authority associated with MC at its original pitch
level is actualized; communicating the story and meaning of the work.’

Appendix H — Prelude’s reprise, counterpoint, child’s transformation — stage 2

d) Prelude’s reprise, counterpoint, child’s transformation — stage 2

Figure 4: Prelude’s reprise, counterpoint, child’s transformation — stage 2

Source: Peter Kaminsky, ‘Of Children, Princesses, Dreams and Isomorphisms: Text–Music Transformations

As the innocent opening statement of the prelude (two oboes) returns in its
complete and unaltered form, the child’s embrace of morality starts to manifest.
Only now, it is grounded in the tonic G major and combined with the second
statement of ‘Il est bon.’
In a metaphoric way, throughout the opera, the morally neutral child progresses in the light of experience and this leads him to become good. There is an isomorphic representation here: the prelude’s tonal neutrality gives way to it confirming to the norms of the tonal system — therefore tonality itself represents moral code.

This example also demonstrates the close music–text relationship in *L’Enfant* and the collaboration between Ravel and Colette.

Appendix I — Continuation of stage 2 and the completion of the child’s transformation at ‘*Maman!*’

e) Maman’s cadence, child’s embrace, final transformation

![Figure 5: Continuation of stage 2 and the completion of the child’s transformation at ‘Maman!’](image)

Bibliography


A Rock Amidst Turbulent Waters: China’s Resilient Centre and its Machiavellian Management of Protests and Petitions

Sam Osborne

Abstract

In contemporary discourse, many struggle to fathom the ability of the Chinese leadership to maintain both economic and social stability despite clear and growing discontent within the Chinese populace. This is not mere good luck, but, rather, good management. Instead of crumbling in the face of widespread protest, as have many other authoritarian regimes, the Chinese leadership utilises Machiavellian methods of managing public discontent to legitimate or strengthen its own rule. In this sense, protest and the ways in which the leadership deals with it are, paradoxically, a driving force behind its overall stability. This essay examines four methods used by the Chinese leadership to respond to protests or petitions, including: the decentralisation of state power; the repression of certain protests and manipulation of facts; the tolerance of and concessions to certain protestors; as well as the cooptation of particularly vehement dissenters. It analyses the ways in which these methods interrelate and how they legitimise the Chinese leadership, ensuring continued stability and the maintenance of China’s rise.

Introduction

Like the steadfast rock within the flowing stream that is so often contemplated by China’s bygone poets, the leadership of the contemporary Chinese Party-state — the Centre — manages to remain resolute and resilient amidst turbulent tides of public discontent. Normally, such unrest might signal the impending demise of a regime, but China’s authoritarian rule shows little sign of decline. This paradoxical ‘coexistence of social disturbance and political stability’ (Cai 2008a: 412) is explained by the Centre’s numerous, rather Machiavellian, methods and political strategies of managing protests and petitions in China,
which, rather than undermining its rule, actually strengthen and legitimises it. In this sense the term ‘Machiavellian’ is utilised to convey methods of statecraft characterised by cunning, deceit and expediency.

The decentralisation of state power and resultant multilevel power structure within China allows the Centre to manage protests in a manner that legitimises the overarching regime. Similarly, repression of certain protests (even violently undertaken) can be manipulated by the Centre to strengthen its own rule. State tolerance of and concessions to other protests and petitions also legitimates the regime in adhering to its propagated legalistic culture. Finally, cooptation by the Centre both of more meddlesome protestors themselves and policies espoused by such petitioners further strengthens its authority. This essay seeks to investigate these ways of managing discontent by the Centre, how they interrelate, and how it legitimises its authoritarian rule — in other words, how the rock manipulates the turbulent waters to cement its own place within the stream. While this essay focuses primarily on the rural population and their protests, it is important to note that urbanites have different perceptions of the Centre and employ different means (particularly involving the Internet) to voice their concerns, continuing the trend of a longstanding ideological and technological divide between rural and urban China.

**Decentralisation of State Power**

The decentralisation of the Party-state’s power in post-Mao China has resulted in a political system with many levels of authority. This allows the Centre to tease out space in the resultant divided power structure to avoid blame and apportion responsibility for dealing with protests and petitions to lower level authorities (Cai 2008a: 411, 415). This structure allows the Centre to fuel its legitimacy and strengthen its rule through such protests by simply ‘[kicking] the ball back into the local government’s court’ (Wu 2009: 200). The Chinese state stands then as ‘less a monolith than a hodgepodge of disparate actors’ (O’Brien & Li 2004: 93–94).

Because of the significant autonomy granted to local authorities under this power structure, there are wide gaps between what the Centre promises and what those local authorities deliver (Chen 2009: 22). This only serves to legitimise the rule of the Centre further, as protests often demand only the removal of unpopular lower level officials, or the local implementation of policies promised by the Centre, rarely questioning the leadership or the Party (Perry 2008: 45).

Due to its concern with protecting the regime’s legitimacy and longevity, the Centre is likely to be (in fact and in the eyes of the people) more tolerant, whereas local officials, concerned as they are with task fulfilment and local issues, are
more likely to be less indulgent (Cai 2008a: 417). The Centre’s ability to direct responsibility for protests to such lower level officials under the divided power structure further strengthens the people’s ‘faith in the good intentions of Beijing’ (O’Brien & Li 2004: 86). This even leads to the belief of many protestors that, in fact, ‘the Centre is on our side’ (O’Brien & Li 2005: 244–45).

The Centre’s management of protests through the decentralisation of state power is perhaps best reflected in a maxim popular in the Chinese countryside:

[T]he Centre is our benefactor (enren), the province is our relative, the county is a good person, the township is an evil person, and the village is our enemy. (O’Brien & Li 1995: 778)

Interestingly, in the literature, commentators such as O’Brien and Li (2005: 240–41) argue that these and other consequences of the multi-layered state in China (such as the Centre being kept ‘in the dark’ and ‘unable to detect breaches of its policies’) are necessarily unwanted by and detrimental to the Centre and its rule.

It is arguable, however, that such commentators fail to adequately appreciate and recognise that the decentralisation of state power, while certainly being heavily flawed, allows the Centre to maintain its legitimacy amidst popular discontent, something thoroughly invaluable to the continued resilience of an authoritarian regime. While this view might be cynical, the alternative interpretation glosses over the Machiavellian nature of the Centre’s management of protests in that the Centre may not desire to ‘break through the host of “protective umbrellas” (baohu san) that local leaders use to fend off oversight’ (O’Brien & Li 2005: 241). In fact, as long as those baohu san remain in place, the Centre can ‘pretend’ not to know of any wrongdoing and escape blame and responsibility, either for the causes of discontent or for the handling of protests by local authorities (Cai 2008a: 430).

Arguably, by placing responsibility for the handling of such discontent in the hands of autonomous, powerful and somewhat impatient authorities keen to dispel the need for Central intervention, the Centre also prevents a ‘lack of discipline’ of protestors and ensures that any popular resistance is ‘short lived, often isolated and does not challenge the regime’ (Cai 2008a: 430), all the while maintaining and even strengthening its own rule.

Clearly, then, by manipulating the overarching divided power structure resulting from the decentralisation of state power, the Centre can safeguard the resilience of its rule by championing the politico-cultural tradition in China that ‘while the local officials are corrupt, the Emperor is always good’ (Zou 2009: 132).
Repression

It might seem paradoxical to consider that repressive conduct by the Party-state in response to certain protests (even violently undertaken) could be managed by the Centre to increase its legitimacy. Yongshun Cai (2008a: 413, 418) argues that such repression undermines an authoritarian regime’s legitimacy and incurs other costs, including risks of retaliatory resistance. While this might generally be so, arguably Cai fails to acknowledge the Chinese Party-state’s sway within the media and its attempts to monopolise all forms of communication. Whereas in contemporary China, the Internet and popular forms of microblogging (weibo) threaten to unsettle such a government monopoly, the Centre still expends tremendous resources in an attempt to control channels of communication and other social media. Through this vast expenditure of both financial and human capital, the Centre can manipulate the way in which protests are framed, and the repression used against them, so as to inevitably favour the state, justify its actions, and strengthen the Centre’s rule. In truth attempts to reframe such incidents are not always successful. The capacity certainly does exist, however, to refashion an overtly repressive response into one that portrays the Centre in a more positive manner.

Of course, if repression of protests occurs locally and is conducted by lower level authorities, the Centre can rely upon the divided power structure discussed above to divert blame and responsibility elsewhere (Cai 2008a: 429). If large-scale repression of protests occurs above those levels, the Centre may still be able to manipulate the petitioning to its own advantage. For instance, after the Tiananmen Square massacre of 1989, suppression of protests was justified by the government’s high-profile propaganda campaign, which was framed to quell ‘alleged “hooligans” and “thugs”’. The Centre reassured its people that ‘the sword of the state remained powerful, swift, and sure’ (Baum 1992: 492).

 Needless to say, any use of violence by citizens in protests allegedly justifies repression by the state (Cai 2008a: 426). Likewise, if the protests threaten social or political stability, even the military can be mobilised (Cai 2008b: 104). This issue of threats to social and political stability invokes the ancient fear held by both governments and citizens alike in China of chaos (luan). By cleverly exploiting that fear and framing protests in terms of affronts to stability, the Centre can inevitably justify any repressive action and, again, strengthen its own rule.

The Centre capitalises on the fear, held by many villagers especially, that any change would ‘unseat the only entity that appears to have the commitment, power, and economic resources to improve their daily circumstances’ (Wright 2010: 8), and also the pertinent worry that a freer society might itself lead to chaos (McCormick 2009: 109). For instance, mass protests erupted at Pubugou,
Hanyuan County, in 2004 over plans to forcibly evict tens of thousands of local farmers prior to planned flooding for the Pubugou Dam. Once the discontent was framed as an ‘illegal act’ and the omnipresent fear of ‘chaos’ was invoked, the protests failed, and any repression by the state would likely have been justified in the eyes of people not privy to the petitions (Mertha 2008: 81).

Furthermore, LHM Ling (1994: 393) provides a fascinating argument of how violent repression of dissent by the state can be warranted under the Confucian discourse of parental governance, and can thereby legitimise and strengthen the rule of the Centre. Under that discourse, ‘state-society relations reflect Confucian parent-child relations’, whereby there exists a natural moral authority on the part of the state as the benevolent but firm parent, versus citizens who are deemed filial dependents (Ling 1994: 393). As a result, the parent-state is granted a license to ‘exile dissenters as ungrateful spoilt brats who deserve to be “rectified”’, and should exercise that moral duty to ‘restore order and harmony in the Confucian family-state’ (Ling 1994: 393–94).

For example, after the Tiananmen Square massacre, Deng Xiaoping dismissed the protestors as ‘dregs of society’ whose real purpose was to ‘subvert our country and subvert our Party’ (Ling 1994: 401).

Further, within this framework, moral suasion as used by the dissenters not only implicitly accepts the ‘hegemonic norms’ that maintain the ruling regime and favour its authority, but are also easily countered by the moral outrage of the parent-state (Ling 1994: 402). For these reasons, then, it is clear that the Centre can manipulate a repressive response to protests in numerous ways, all of which (paradoxically) strengthen and legitimise its own rule.

Tolerance and Concessions

The Centre also uses tolerance and concessions as means of managing protests and petitioning. Interestingly, Cai (2008a: 413, 414) argues that concessions are detrimental to the legitimacy of an authoritarian regime, as is tolerance, which he contends will lead only to more resistance.

Again, it is arguable that Cai fails to situate tolerance and concessions within the context of legitimacy derived from continued economic growth, as well as the Confucian context of modern Chinese political culture, in which such acts will only further add to the ‘benevolent’ and ‘forgiving’ image of the parent-state, strengthening its rule. Regardless, the Centre is more likely to be tolerant and to concede to demands of protestors when discontent is pursued through the legal channels that are provided for by the state. By fostering a culture of alleged legalism, which mandates strict obedience to the legal system and an
overarching rule of law, the Centre thus manages to artfully manipulate such protests and petitions to its own advantage. If dissatisfaction is aired through legal avenues or employs legal norms imposed by the Centre, this only further legitimises their rule.

The Centre has provided various lawful channels through which dissenters can voice their grievances, in doing so ensuring a ‘win-win’ situation with regard to their legitimacy: ‘policy-based resisters’ (diaomin) utilise ‘laws, policies and other official communications to defend their interests’, thereby simultaneously attesting to the legitimacy of the Centre’s rule. Conversely, ‘recalcitrants’ (dingzihu), or ‘illegal’ protestors, invoke the Centre’s inherent right to discipline wrongdoers and can be dealt with under law, thereby, again, confirming the Party-state’s rule (Li & O’Brien 1996: 31, 35).

An example of a lawful method provided for by the Centre through which a protestor can voice dissatisfaction is the Administrative Litigation Law introduced in 1989. While under this law citizens can sue malfeasant local officials, the Party retains immunity from suit (O’Brien & Li 2004: 80). Again, this means that the Centre manages to further exploit the divided power structure of the Party-state by lawfully allowing protestors to target lower level authorities while retaining higher level regime validity.

Further, the Centre has given its citizens ‘more laws (and other directives) to point to when filing charges’ (O’Brien & Li 1995: 764), fuelling the people’s belief in the Centre’s alleged adherence to legalism and further increasing its legitimacy when such laws are then utilised and responded to, often with tolerance or concessions.

Another lawful channel of protest managed by the Centre, and one that is again likely to receive responses of forbearance or indulgence, is the xinfang (or ‘people’s letters and visits’) system of collective petitioning. While this is an additional legal avenue for citizens to ‘air their grievances’, it is also a means by which the Centre can ensure a ‘close connection with the masses’ (Zou 2009: 124, 126). Agreeably, Keyuan Zou (2009: 132, 134) goes as far as to say that it is ‘a means of controlling society and the people’, and that, for the Centre, the ‘ideal model of the xinfang in its mind would be like “a bird in a cage”’ — in other words, in line with the reasoning put forward above, a lawful means of protest, highly controlled by the Centre, which is inherently rule-legitimating. Moreover, by tolerating and even condoning such contention, this arguably also solves the principal-agent problems resulting from the decentralisation of power (which, as shown above, are simultaneously manipulated by the Centre to its own advantage) (O’Brien & Li 2005: 241).
Thus the Centre manages protests in this way by ‘empowering ordinary people to serve as watchdogs’ (O’Brien 2003: 60), strengthening its rule by giving its citizens a sense of self-worth while concomitantly ‘getting local feedback without holding substantive elections’ (Weller 2008: 122).

Moreover, those diaomin (or ‘rightful resisters’ as O’Brien and Li dub them) who rely upon laws and policies promulgated by the Centre to frame their contention situate themselves within the culture of legalism championed by the Party-state, thereby strengthening its rule further. Their resistance is generally ‘mindful and circumscribed, local and parochial, rather than national and autonomous’ (Li & O’Brien 1996: 54), and thus the Centre is more likely to react with tolerance or concessions rather than violent repression (although this is dependent on the media’s input), which only reinforces their ‘benevolent’ rule.

In addition, Chinese politico-cultural tradition subscribes to a general ‘right to rebel’ stemming from Mencian thought (Perry 2008: 39) and, as diaomin also endorse ‘waving the red flag to oppose the red flag’ (Perry 2002: xxiii), they inevitably legitimise the pre-existing authoritarian rule. Any rights such resisters call for are also ‘granted, not recognized’ by the Centre (Hintzen 1994: 12–13), and thus advocating the implementation of such rights in turn presupposes the legitimacy of the Centre.

Clearly, then, by responding with tolerance of or concessions to those protests or petitions that operate within the Party-state’s sphere of legality, the Centre is able to manipulate such discontent so as to strengthen its own legitimacy and rule, fortifying the tradition of ‘critique within the hegemony’ (O’Brien & Li 2006: 5).

Cooptation

Finally, the Centre can also respond to protests by either coopting more meddlesome petitioners themselves into Party-state political mores, or by adopting such dissenters’ policies. While both coopting protestors themselves and ‘implementing the dissenters’ reforms (Ling 1994: 402) potentially jeopardises the Centre’s legitimacy as signs of weakness, again, this argument fails to hold water when the Confucian context of parental governance in China is taken into account. Here, such actions only strengthen the ‘benevolent’ rule of the Centre.

In fact, as Elizabeth Perry (2008: 45–46) rightly argues, the Centre’s willingness to respond in such ways shows flexibility and pragmatism, an essential feature of Chinese morality and state behaviour. This is invaluable in ensuring the longevity of their rule. Furthermore, incorporating more troublesome individuals into the existing power structure enables the Centre to find favour.
with those who could potentially upset its rule (including outsiders), as well as displaying forbearance and acceptance to its citizenry. An example is the Party’s cooptation of the outspoken Yu Jianrong as a kind of teacher of ethics to state officials (Li 2010). The forthright Yu, who uses social media to buttress his campaign to prevent abuse of petitioners, was placed under a government gagging order after publicly contradicting propaganda suggesting that the rural petition system had achieved ‘groundbreaking success’ (Li 2010). Nowadays, however, Yu is regularly employed to lecture state officials on ethics and other moral guidelines in relation to the handling of protests and petitions, which exemplifies how the Centre manages to strengthen its rule both by keeping disruptive individuals firmly within its sphere of influence, as well as instilling its own agents with values less likely to instigate discontent in the first place.

Conclusion

Ultimately, then, the Centre uses various means of managing protests and petitions within China in an often Machiavellian manner to undergird and strengthen its own rule. It has been able to manipulate the decentralisation of state power and its resultant multilevel power structure to allow blame and responsibility for protests to be apportioned to lower level authorities, escaping accountability and maintaining its legitimacy as the ‘benefactor’ of the Chinese people. Moreover, even where violent repression is used in response to petitioning, the Centre is generally able to justify any such repression under Confucian mores, in turn strengthening its place at the top of the political hierarchy. Tolerance of and concessions to protests are more likely where dissenters utilise legal frameworks of contention provided for by the state, which inevitably presuppose and perpetuate the validity of the Centre’s rule. Cooptation of particularly irksome protestors, or the reforms advocated by such protestors, also further reinforces the supremacy of the Centre’s ‘benevolent’ rule.

Although protestors, like water in the stream, are indeed a strong force that might smooth the edges of a rock over time, inevitably that rock, like the Centre, can manipulate those waters and cement itself firmly within the stream for the foreseeable future.
Bibliography


Was there a Political Dimension to the Irish Rockite Movement of 1821 to 1824?

Rebecca Preston

Abstract

The Rockite movement was one of the most violent agrarian rebellions in Irish history, renowned for its use of incendiarism and murder to achieve a basic form of political mobilisation. The Rockite movement, along with other pre-famine agrarian rebellions, was traditionally perceived to be motivated by purely localised socio-economic tensions. Through analysing the famous Rockite notices and contemporary newspaper commentary of Rockite attacks, it is evident that the Rockites were a complex and unique movement, transcending more orthodox understandings of social banditry. The Rockites were motivated by a strong sense of political agitation and significantly influenced contemporary, as well as post-famine politics. The violent and conspiratorial tradition of the Rockites continues to resonate with modern paramilitary organisations in Ireland.

The Irish Rockite movement of 1821 to 1824 was characterised by sustained agrarian violence and the use of intimidating documents by Irish agrarian rebels. The Rockites were named after their mythical leader Captain Rock and dominated southern Ireland throughout the early nineteenth century. Due to Daniel O’Connell’s high profile constitutional movement for Catholic emancipation throughout 1823 to 1847, pre-famine agrarian rebellions have traditionally been viewed as apolitical and led by lower class social bandits who were motivated purely by localised socio-economic tensions. SJ Connolly, for example, argued that the Rockite movement was ‘primarily a pragmatic, even conservative, movement, concerned with limited and specific’ economic-based goals, including the regulation of rents, wages and tithes, the protection of poor tenants threatened with eviction and wider access to land for tillage.1 Connolly dismissed the notion that the Rockites were motivated by political agitation, concluding that the Rockite movement was ‘not radically different from other

---

agrarian protests of the early nineteenth century’. 2 Eric Hobsbawm similarly contributed to the romanticised conception of Irish agrarian rebels as merely lower class bandits agitating for improved economic circumstances. In his book *Bandits*, Hobsbawm viewed social banditry as a common trait inherent within all rural societies in ‘old-established and permanent pre-industrial empires’. He argued that such societies, including Ireland, have long developed traditions of ‘unofficial political systems and networks’, which often reached ‘out to all who were outside and against the official structure of power, including both social bandits and the outsider groups’. He claimed that only those organisations that had strong community alliances and resources could be described as effective political rebels, rather than mere social bandits. 3 To an extent, the Rockites conform to Hobsbawm’s conception of banditry, as members of the Rockite movement were often poor outsiders refusing to ‘accept the normal rules of poverty’ and in opposition ‘to the hierarchy of power, wealth and influence’ of Ireland. 4 Conversely, the Rockites challenge Hobsbawm’s homogenised definition of social bandits, as the complex movement was supported by a multiplicity of social classes and was motivated by traditional economic grievances as well as political and religious factors. The Rockite movement demonstrates that the distinction between social bandit and political leader is often blurred.

This essay challenges the orthodox view put forward by Connolly and Hobsbawm, by proposing that the Rockites were significantly influenced by contemporary politics and played a central role in pre-Famine Irish political life. 5 It argues that the Rockite movement had a strong political dimension as they were perceived as a political threat by the British Government and were partially motivated by political grievances. It is acknowledged that the Rockites were not solely motivated by political agitation — the movement encompassed a multifaceted agenda. The myriad motivations, however, including economic, ideological, political and religious, were interconnected and contributed to the politicisation of the Rockites. This essay aligns more closely with the arguments proposed by Tom Garvin, Shunsuke Katsuta and James Donnelly. Although focusing on the secret society of the Ribbonmen, Garvin highlights the important political role secret societies played in pre-Famine Ireland, arguing that the Ribbonmen ‘showed distinct signs of politicisation’ and formed an important ‘connecting link with the revolutionary nationalist societies’ of the eighteenth century. 6 Katsuta engaged in a localised study of the Rockites, focusing on the political, millenarian and agrarian dimensions of the Rockite

---

2 ibid., 82.
5 The term ‘political’ is used in reference to systems of government, state policies, and ideologies relating to state power.
movement in County Cork.\footnote{Shunsuke Katsuta, ‘The Rockite Movement in County Cork in the Early 1820s’, \textit{Irish Historical Studies}, 33 (2003), 279.} Through his numerous articles and books, Donnelly provides the most comprehensive historiographical contribution to the history of the Rockites. Donnelly particularly stressed the sectarian and millenarian dimensions of the Rockite movement, arguing that Pastorini’s prophecies (expounded in Bishop Charles Walmesley’s book \textit{The General History of the Christian Church}, published under the pen name of Signor Pastorini\footnote{James S. Donnelly, Jr., \textit{Captain Rock: The Irish Agrarian Rebellion of 1821 – 1824} (Madison, 2009), 122.}) ‘helped to rally Catholic country people to the Rockite cause’.\footnote{ibid., 10.} He similarly recognised the political features of the Rockite movement, but cautions that he does not believe the ‘Rockite movement was predominantly political’.\footnote{ibid., 22.} Unlike other historical works, this essay exclusively analyses the Rockite movement from a political perspective. By doing so, it is evident that the actions and motivations of the Rockites significantly influenced the political history of post-Famine Ireland, particularly the nationalist and sectarian violence which plagued twentieth century popular politics. The violent and conspiratorial tradition of secret societies continues to resonate with modern paramilitary organisations, such as the Irish Republican Army (IRA) and Ulster Volunteer Force (UVF), which are still active in Northern Ireland today.

This essay begins with a brief overview of the secret societies in Ireland, the origins of the Rockites and the economic circumstances which significantly influenced the Rockite movement. The political dimension of the Rockite movement is demonstrated firstly through analysing the nationalist and anti-English language of Rockite notices and the promulgation of Rockite regulations. Secondly, this essay assesses the Rockites’ use of symbolic violence to target Protestant clergymen, particularly those representing the Church of England, and representatives of the British state. It also analyses British responses to such violence through the introduction of specialised British laws and the army. Finally, this essay recognises, as Donnelly stresses, the importance of millennialism and sectarianism to the Rockite movement, but goes further to propose that the religious influences of millennialism and sectarianism were largely politicised, as the Rockites viewed Protestantism and the British state as the ultimate cause of Catholic poverty and oppression. Contemporary newspapers, predominantly \textit{The Times}, \textit{Dublin Evening Post} and the \textit{Morning Chronicle} are drawn on to provide evidence of political motivations and details of Rockite actions, as newspapers frequently published Rockite notices.

The phenomenon of secret societies was a peculiarly Irish response to social and political grievances. Secret societies flourished in rural Ireland, particularly in the southern counties, between 1760 and 1845 when they came to be central
to pre-famine political life in Ireland. Particularly in the 1820s, the height of the Rockite movement, Irish popular politics was significantly influenced by violent underground movements. These multifaceted societies had their own agenda, motivations, tactics and class composition. The Caravats, for example, were a unified rebel group of lower class rural labourers and farmers, whilst the Shanavests were primarily of middle-class origins. Common to all societies was the use of agitation, intimidation and violence by agrarian rebels to achieve a basic form of political mobilisation. The Rockite movement constituted one of the most extensive and violent agrarian rebellions in pre-famine Ireland, renowned for its frequent use of incendiarism and murder as weapons of warfare.

The Rockite movement began in the west of County Limerick in 1821, following the assassination of the son of Alexander Hoskin — the land agent of the Viscount Courtenay’s 34,000-acre estates. Hostilities towards Hoskin were accentuated after he increased the tenants’ annual rent by 40 per cent. Following the assassination, small groups of rebels began to raid towns in counties Cork, Limerick and Kerry, acquiring arms and ammunition. Dublin Castle, which served as the administrative centre for the British Government, army and police, immediately dispatched police forces, troops, yeomanry corps and veteran battalions to repress the escalating disturbances. These measures forced the first Captain Rock, leader and assassinator of Hoskin’s son, Patrick Dillane, to flee into undeveloped rural land. Dillane was a blacksmith from the Shanagolden district and was a tenant of Lord Courtenay’s estate during Hoskin’s tenure as agent.

From 1821 onwards there is evidence of secret Rockite societies forming across the province of Munster through the swearing of oaths. Similar to other secret societies such as the United Men, Ribbonmen and the Caravats, members of the Rockites were sworn by an initiation oath which contained vows of sectarian and political nature. The vows were intended to safeguard against betrayal by forcing Rockite members to swear allegiance to the movement and to assist in destroying all non-Catholic heretics and oppressors. Despite the common Rockite oath, it remains unclear to what extent the various Rockite groups across Munster were coordinated and organised as a common underground organisation.

---

12 Garvin, ‘Defenders, Ribbonmen and Others’, 133.
17 Donnelly, Captain Rock, 97.
18 ibid.
A unique feature of the Rockite movement was the complexity and diversity of its social composition and the widespread support it received from all classes in rural Ireland. Both wealthier, middle-class farmers, as well as poorer farmers, labourers, artisans and cottiers from southern and south-west Ireland, were represented among Rockite supporters and leaders. This broad social composition of the movement was unique among agrarian rebellions. Donnelly argues that ‘there were certain grievances that promoted social cohesion across class lines below the elite, just as there were other grievances that gave rise to acute conflict between one non-elite social group and another’. Thus, the level of social cohesion or disunity between Rockite groups fluctuated depending on the economic conditions and nature of the landlord–tenant relations in a particular district. This enhanced cross-class collaboration was a product of the Rockites’ alliance with Pastorini’s millenarian prophecies, the use of sensational violence and contemporary economic depression, which resulted in tithe and rent grievances being pushed up the social scale. The intensity of the economic depression between 1821 and 1824 brought substantial economic hardship to once prosperous farmers and graziers. Donnelly suggests that this ‘burning sense of economic injustice’ primarily accounts ‘for the diverse social composition of the Rockite movement’.

The contemporary economic circumstances influenced the Rockite movement by engendering political agitation, millennialism and sectarianism. The series of economic crises in early nineteenth century Ireland aroused widespread discontent among rural society, enabling the Rockite movement to thrive and grow to maturity. The primary grievances of the Rockites were a product of the economic and agrarian climate of the early nineteenth century. The conversion of grassland, which was largely or completely tithe free, to sites of grain production, resulted in increases in taxes. In Munster and Leinster, tithes were charged on potato crops at the highest rates per acre. These areas coincided with a greater prevalence of Rockite violence. During 1819 to 1823, Ireland experienced a collapse in agricultural prices, particularly grain and livestock. This was partly due to the fact that the Irish economy was almost exclusively dependent on the British market, resulting in full exposure to British post-Napoleonic war price deflation and fluctuation. Donnelly noted that the plunging prices of grain and livestock alone were sufficient to incite widespread agrarian rebellion. This economic depression was exacerbated by the fever epidemic of 1816 to 1819 and the partial famine of 1821 to 1822 — both of which aroused fear and desperation.

---

19 Donnelly, Captain Rock, 341.
20 ibid., 150.
21 ibid., 151.
23 ibid.
24 ibid., 31.
25 Donnelly, Captain Rock, 340.
among rural Ireland’s population. 26 The economic crises of the early nineteenth century enhanced cross-class collaboration between the landless and the landed and fostered widespread support for central Rockite grievances, particularly tithes, exorbitant rents, evictions and escalating food prices. 27 The severity of the economic depression meant that discontent about rents and tithes were strongly felt by the wealthier farmers and graziers. Thus, agitation to regulate tithes and rents, and control the occupation of land became a primary concern of both poor and wealthy farmers. The Rockites’ use of sensational violence further appealed to the masses.

The Rockites championed their interests through humiliating notices, threatening letters and violent attacks on property, predominantly through incendiary methods, and individuals, often beating, mutilating or murdering their victims. These extreme techniques were intended to display Rockite authority and incite fear throughout Ireland. Rockite violence was underscored by strong political symbolism, primarily through targeting representatives of the existing political system, including the Protestant clergy and the British army. By 1824 the movement had drastically dissipated, primarily as a consequence of the strong measures taken by Dublin Castle and the British Government.

The anti-English and nationalist sentiments espoused in Rockite notices demonstrate the influence of contemporary politics on the Rockite movement. In the first place, it is significant that the Irish-speaking peasantry used the English language in Rockite notices. The Rockites’ use of English vocabulary was a political technique borrowed from the agrarian rebels of the 1790s. 28 The Rockites expressed profound animosity towards British subjugation and political injustice to justify extreme violence. Evidence of anti-English resentment is demonstrated in a notice posted in County Tipperary:

We are pressed by the tyrannical laws of the English government … to which they crowd us up with rent, tythes, and taxes … English laws must be curbed in, for we will never be satisfied until we have the Irish parliament and king crowned in Ireland. 29

The provocative and emotive language of this Rockite notice arouses anti-English sentiments and exemplifies the ways in which economic grievances were politicised. The use of ‘we’ positions the Rockites in direct opposition to the English, depicting themselves as a united, revolutionary force. This insurrectionary language was similarly employed in a notice posted outside Adare chapel in Limerick, which ordered the mobilisation of men ‘in the name of the

26 ibid., 25, 340.
27 ibid., 341.
29 Dublin Evening Post, 27 Dec. 1821.
Was there a Political Dimension to the Irish Rockite Movement of 1821 to 1824?

Irish’ to march ‘against the common enemy, the tyrant of Ireland, the English’ in order to gain Irish independence. The juxtaposition between English tyranny and Irish freedom serves to create a sense of solidarity among the Rockites and legitimise their movement as representative of the Irish population. The explicit nationalist language revives the vision of the 1790s revolutionary movement. The anti-English and nationalist sentiments of Rockite notices were manifested through the articulation of their own standards of justice.

By promulgating their own laws and demanding strict adherence, the Rockites directly challenged the political authority of the Irish Government and the British Empire. After County Cork was declared under the operation of the Insurrection Act 1822, the Rockites sought to undermine the legal legitimacy of the state by posting a ‘violent and inflammatory’ notice on the chapel at Cloghroe. The notice proclaimed that inhabitants of County Cork should not pay tithes or taxes and all land taken within seven years must be surrendered. The Times reported that ‘insurgents’ attacked a ‘respectable farmer’ as he failed to surrender his farm in accordance with the notice. By promulgating laws and exacting punishment on lawbreakers, the Rockites mimicked and thereby challenged the established legal and judicial system. Rockite punishment often involved death or property destruction by incineration. A notice posted near Shambles, ordered that ‘any person taking the contract for Soldiers’ meat … shall be met with fire and destruction’. By targeting soldiers — representatives of the state — the Rockites imposed their political agenda on rural society, using intimidation to ensure their desires were achieved. The threat of violence consolidated Rockite power through manipulating the psychology of fear. This is highlighted in a notice sent to Mr Lamvert, the proprietor of the Kilnapp Iron Mills. Captain Rock threatened that if Lamvert failed to ‘obey’ his orders, he will be ‘consumed to ashes and if not, you or any of your family shall be watched … until you share the same fate which others have shared.’ Like political authorities, the Rockites enforced harsh punishments and used their victims as examples.

The notice sent to Lamvert was also employed to champion a nationalist agenda, evidenced by Captain Rock’s signature — ‘[signed] your sincere friend the brave and powerful Commander of the United Forces of Ireland’. By identifying Captain Rock as a ‘Commander of the United Forces of Ireland’, the

30 Dublin Evening Post, 16 Feb. 1822.
32 The Times, 4 Mar. 1822.
33 ibid.
34 ibid.
35 Morning Chronicle, 9 Jan. 1822.
36 Dublin Correspondent, 22 May. 1823.
37 Donnelly, Captain Rock, 95.
38 Dublin Correspondent, 22 May. 1823.
Rockites revived the language and vision of the revolutionary movement of the 1790s and created an image of a unified movement mobilised against political authorities. This image of the noble vigilante, addressing the grievances of the poor, was important in fostering widespread support and adding cohesion to the movement. This is illustrated in Captain Rock’s letters to James Condon, ordering him to stop “tyronizing over poor people … and leave poor people their potatoes and turf”.39 Through establishing their own regulations and advocating as a united force on behalf of the Irish poor, the Rockites presented a significant political challenge against British authorities.

Rockite violence was underscored by strong political symbolism, mainly by targeting representatives of the existing political system. Frequent violent attacks on property and individuals resulted in the mobilisation of all available resources by Dublin Castle to contain the Rockites. Special legislation was enacted to regulate rural disorder and political agitation. To restrict the use of intimidation and punishment by the Rockites and other secret societies, the British Government enacted the *Whiteboy Amendment Act 1831*40 and the Insurrection Act, which suspended habeas corpus and imposed a night curfew.41 Habeas corpus is a legal writ which requires an arrested person to be brought before a judge or jury in court, essentially protecting arrested persons from unlawful detention. The suspension of this writ resulted in trial without jury. The Insurrection Act was introduced in eight Irish counties; in Munster alone over 1500 men were trialled, of whom 200 were convicted and transported.42 The *Unlawful Oaths (Ireland) Act 1823* and the *Party Processions (Ireland) Bill*43 were designed to repress sectarian violence, targeting both Catholic agrarian rebels and Orange societies. The Orange societies were a decentralised political organisation named after the Protestant William of Orange, King William III of Great Britain. In response to growing agitation for Catholic emancipation, the Orange Order, originally named the Orange Society, was formed in 1795 to maintain Protestant ascendancy.44 The introduction of British laws directly targeting agrarian violence highlights the direct political challenge the Rockites posed to the British state and British–Irish relations.

39 *Morning Chronicle*, 1 Dec. 1821.
40 The original title of the Act referred to is the *Tumultuous Risings (Ireland) Act 1831*, which is an amendment of the *Tumultuous Risings Act 1775*, commonly known as the *Whiteboy Act (15 & 16 Geo. 3. c. 21)*.
44 The Orange societies continue to exist in Ireland and remain closely associated with Ulster Unionism.
Rockite violence was also suppressed through the deployment of five British regiments, yeomanry and police.\textsuperscript{45} Lord Sidmouth, when acting as home secretary, admitted that the British army was deployed in southern Ireland to put down a ‘most formidable Danger’.\textsuperscript{46} The introduction of armed forces into southern Ireland exacerbated sectarian animosities and intensified campaigns of violence.\textsuperscript{47} The Rockites often targeted the yeomanry and British army as they represented Protestant ascendency and imperial subjugation — the perceived sources of Catholic destitution and economic oppression. Rockite attacks against the British military were most formidable in County Cork, with incendiaryism occurring at a level ‘unprecedented in the history of Irish agrarian disturbances’.\textsuperscript{48} It is noteworthy that the Rockites did not exclusively target British and Protestant victims. Gale Christianson records that, despite the ‘long-standing enmity between Catholics and Protestants’, many agrarian rebels ‘cut across the boundaries established by religious conviction’ if it was necessary to achieve an ultimate goal, for example ending the tithe system.\textsuperscript{49} In addition, the Catholic clergy strongly condemned the tactics and violence employed by agrarian rebels — further exacerbating hostilities between Catholic peasantry and the Catholic Church.\textsuperscript{50} Yet overall, the Rockites primarily directed violence towards Protestant authorities and such attacks were underscored by strong political motivations. The yeomanry deployed in southern Ireland were associated with Protestant triumphalism as they participated in the Orange celebrations of the anniversary of the Battle of the Boyne — a politically charged event symbolising the beginning of British imperialism, Protestant ascendancy and the destitution of Irish Catholics.\textsuperscript{51}

To reduce the risk of detection, the Rockites offered rewards to strangers to carry out violence, evidenced by the letter received by a sergeant of the Adare yeomanry. Captain Rock offered ‘twenty pounds for every [yeoman] put to death; one hundred pounds for every officer; and fifty pounds for a serjeant [sic]’.\textsuperscript{52} By attacking representatives of the state, the Rockite movement became intrinsically political. Furthermore, the Rockites frequently exacted violent punishments whilst dressed as women, referring to themselves as ‘Lady Rock’.\textsuperscript{53} A similar phenomenon occurred in other secret societies, with the Whiteboys producing Lady Clares and Molly Maguires.\textsuperscript{54} By reversing gender identities,

\begin{itemize}
\item \textsuperscript{45} Donnelly, \textit{Captain Rock}, 344.
\item \textsuperscript{46} Lord Sidmouth, 23 October 1821, quoted in Katsuta, ‘The Rockite Movement in County Cork’, 278.
\item \textsuperscript{47} Donnelly, \textit{Captain Rock}, 268.
\item \textsuperscript{48} Katsuta, ‘The Rockite Movement in County Cork’, 278
\item \textsuperscript{49} Gale E Christianson, ‘Secret Societies and Agrarian Violence in Ireland, 1790–1840’, \textit{Agricultural History}, 46 (1972), 375.
\item \textsuperscript{50} ibid.
\item \textsuperscript{51} Donnelly, \textit{Captain Rock}, 139.
\item \textsuperscript{52} Caledonian Mercury, 8 Nov. 1821.
\item \textsuperscript{53} Dublin Evening Post, 5 Feb. 1822.
\item \textsuperscript{54} Donnelly, \textit{Captain Rock}, 110.
\end{itemize}
the Rockites inverted social and sexual roles, repositioning women ‘on top’ and transforming women’s traditional role of motherly protector to vicious attacker.\(^{55}\) This gender inversion paralleled the repositioning of the lower class Rockites as political governors and legal authorities of rural Ireland. The cross-dressing could also reflect a broader critique of British–Irish relations, repositioning the feminine Hibernia to rule over masculine Britannia. The Rockites’ use of targeted violence and cross-dressing reflects a rejection of contemporary society and the political system.

The Rockite movement was motivated by sectarianism and Pastorini’s millenarian prophecies, which envisioned the obliteration of Protestant ascendency in Ireland in the year 1825.\(^{56}\) Religious divisions in Ireland were accentuated by political realities, as the legal system secured Protestant monopoly of landownership, government positions and the economy. Pastorini’s *Book of Revelations*, which predicted the end of Catholic destitution through the destruction of Protestant dominance, resonated with rural Irish Catholics across all classes, adding cohesion to the Rockite movement.\(^{57}\) The strong religious and anti-Protestant undertone of the movement reduced the sharpness of class antagonism and enabled the Rockites to rally cohesive support among rural Catholics.\(^{58}\) This millennial upsurge in Ireland was a result of inflated agricultural prices, deep-seated sectarian animosities, the revival of aggressive Protestant evangelicalism in 1815 and the growing intensity of the tithe grievance.\(^{59}\) The exclusion of poor, Catholic peasants from the political apparatus of the state further popularised the millennial movement and led to the idealisation of the Rockites as an alternative political vehicle through which they could advance their common interests.\(^{60}\) The Rockites used millenarian rhetoric to dehumanise the Protestant population, thus allowing the Rockites to morally justify and legitimise the use of extreme violence and incineration against Protestant landlords, clergy and churches. The *Morning Chronicle* reported the destruction of the Anglican church of Knockane which was ‘maliciously set on fire’ as a threat to churchwardens to stop collecting tithes.\(^{61}\) The tithe grievance was central to the Rockites and demonstrates the interconnection of economic, religious and political motivations. The collection of tithes for the Church of Ireland fostered resentment among Catholics and Protestants as it represented British imperialism, enhanced economic destitution of the rural poor and, for

\(^{55}\) ibid., 8.
\(^{56}\) ibid, 52.
\(^{57}\) ibid, 52.
\(^{59}\) ibid., 135–37.
\(^{60}\) Donnelly, *Captain Rock*, 7.
\(^{61}\) *Morning Chronicle*, 4 Dec. 1821.
the Catholics, symbolised Protestant ascendency.62 A letter sent from Captain Rock to a Mr Hume highlights the centrality of tithes to Rockite violence and the politicisation of economic grievances. Joseph Hume (1777-1855) was a Scottish Member of Parliament who sat in the House of Commons. Hume was a strong advocate for free-trade principles and radical social reform.63 Captain Rock targeted Hume as he believed that Mr Hume was ‘getting Parliament to do away with this curse-of-God tithe-law’. Captain Rock explained that ‘house-burnings’ and assassinations of ministers and tithe proctors were the only method of ‘letting Parliament know that the savage tithe-law is galling us’.64 Captain Rock further advocated for the repeal of the Insurrection Act and the restoration of habeas corpus, demonstrating the Rockites’ use of threatening letters and violence as a vehicle of political protest.

The Rockites transcended traditional notions of social banditry and apolitical agrarian rebels. The Rockite movement had a strong political dimension, evident through the nationalist sentiments in Rockite notices and the promulgation of Rockite laws which undermined the legal and political authority of the state. Violence was often used as a vehicle of social and political protest, challenging Protestant and British ascendancy in Ireland. The politicised nature of the Rockites is also evidenced by Britain’s response, namely the introduction of specialised laws and the deployment of armed forces. The frequent publication of Rockite notices in contemporary newspapers and British descriptions of the Rockites as malicious, notorious and savage, suggest that the Rockites were perceived as a substantial political threat.

Bibliography

Primary

Newspapers

_Caledonian Mercury_

_Dublin Correspondent_

_Dublin Evening Post_

_Finns Leinster Journal_

_Freemans Journal_

62 Donnelly, _Captain Rock_, 6.
63 _The Times_, 4 Mar. 1822.
64 ibid.
Morning Chronicle

The Times

Hansard Debates

Hansard Debates of 1832 (HC Deb 25 June 1832 vol 13 cc1023-51).

Secondary

Books


Articles


Christianson, Gale E, ‘Secret Societies and Agrarian Violence in Ireland, 1790–1840’, *Agricultural History*, 46 (1972), 369–84.


Was there a Political Dimension to the Irish Rockite Movement of 1821 to 1824?


Mainstreaming Restorative Justice in South Australia’s Criminal Justice System: A Response to the Over-Representation of Indigenous Offenders

Alexandra Smith

Abstract

John Braithwaite, a leading advocate for restorative justice, stated that: There can be no justice in a world without connectedness and empathy; at the same time, social capital cannot flourish in a world without an infrastructure of security around human relationships that can only be guaranteed by institutions of justice.1

In recent years, the concept of restorative justice has attracted much attention from policy makers, legal practitioners and social justice advocates. This essay discusses the shortcomings of South Australia’s current court system and its failure to adequately respond to the needs of Indigenous offenders, and considers the potential for the increased use of principles of restorative justice to provide beneficial outcomes in addressing those needs. The following is a consideration of how mainstream criminal sentencing can be reimagined to integrate restorative justice, and suggests that South Australia adopt legislation based on the Crimes (Restorative Justice) Act 2004 (ACT) in order to mandate restorative justice considerations as a compulsory part of the criminal sentencing process.

Introduction

Over the last three decades restorative justice has been established as a legitimate response to criminal behaviour.2 The benefits of restorative justice are increasingly being recognised internationally, with the United Nations encouraging states to develop guidelines and standards to govern the use

of restorative justice programs. In line with Braithwaite’s work, this essay argues that the principles of restorative justice can provide the foundations of strong institutions of justice built around the values of connectedness and empathy. In particular, given the High Court’s recent recognition of the need to consider an offender’s Indigenous background in *Bugmy v The Queen*, this essay outlines the extent to which restorative justice may improve South Australia’s criminal justice system to better serve the needs of Indigenous offenders and their communities.

The essay is divided into four parts. The first part considers the historic disadvantage faced by Indigenous peoples, and the inadequacy of South Australia’s current criminal justice system in addressing the underlying causes of offending amongst this group. The second part proposes a solution to these issues in the form of legislation that creates a centralised restorative justice scheme to augment existing legislation. It is not the purpose of this essay to put forward a completed draft of legislation; rather, it outlines a legal framework largely modelled on the *Crimes (Restorative Justice) Act 2004*, which was recently introduced in the Australian Capital Territory. The third part considers the beneficial outcomes of such legislative change and the fourth explores the drawbacks of restorative justice and the political feasibility of the proposed legislative changes. The essay concludes that the SA Government should consider introducing legislation that will implement the principles of restorative justice in order to address the underlying issues that may cause criminal behaviour amongst all offenders; in particular Indigenous offenders.

**Definition**

The most commonly cited definition of restorative justice describes it as ‘a process whereby all the parties with a stake in a particular offence come together to resolve collectively how to deal with the aftermath of the offence and its implications for the future’. Restorative justice processes, however, do not necessarily involve face-to-face meetings between victim and offender. A broader approach to restorative justice recognises that the method encompasses a variety of practices at different stages in the criminal justice process, including conferences between victim and offender, diversion from court prosecution, redirection to rehabilitation programs — such as drug and alcohol therapy or anger management programs. This essay uses the definition of restorative justice outlined in the Explanatory Statement accompanying the ACT’s *Crimes*...
(Restorative Justice) Bill 2004, as a methodological tool to address underlying issues experienced by all those involved in the criminal justice system, including victims, offenders and their communities.6 Most relevant to this essay is the aim of restorative justice to have a constructive impact on the life of the offender.7 This essay focuses on how restorative justice processes may achieve this at the sentencing stage of the criminal justice process.

Part One: Historic Disadvantage and the Inadequacy of Traditional Court Systems

Historic Disadvantage

The imprisonment rate of Aboriginal and Torres Strait Islander peoples in Australia is considerably higher than the rest of the population in all Australian states and territories. South Australia has historically had the highest representation of Indigenous people in prison,8 with recent figures indicating that Indigenous men are 9.5 times more likely to be arrested than non-Indigenous men, and Indigenous women 16.3 times more likely to be arrested than their non-Indigenous counterparts.9 What is perhaps most striking is that this disproportionate representation of Indigenous people in the criminal justice system is unlikely to improve in the future, with young Indigenous people also being significantly over-represented in the criminal justice system. Young Aboriginal and Torres Strait Islander people account for approximately 1.2 per cent of the overall population of young people in Australia;10 however, they make up 21 per cent of young people on community justice orders, and 23 per cent of young people held in detention.11 Furthermore, despite the proven benefits of diversionary processes such as restorative justice conferencing, South Australia has recently been identified as a jurisdiction that is less likely to adopt these methods.12

---

10 Australian Bureau of Statistics ‘Experimental Estimates of Aboriginal and Torres Strait Islander Australians’ (Release no. 3238.0.55.001, June 2006).
In order to understand the significant over-representation of Indigenous people in Australia's criminal justice system, it is necessary to consider the ongoing systemic damage and chronic marginalisation of Indigenous communities. Australians' Indigenous communities have suffered systematic damage in the form of expropriation of their lands and the destruction of their culture and heritage and have endured governance that has effectively removed self-determination from their lives. Indigenous communities have been left emotionally traumatised and socially and economically excluded from mainstream society. Further, Australia's common law recognises that many Indigenous offenders come from backgrounds of social disadvantage characterised by abuse of alcohol and violence. It is well established that social and economic disadvantages engender antisocial and criminal behaviour. Additionally, systematic bias against Indigenous people within the criminal justice system and the impact of over-policing have contributed to this over-representation.

The need for courts to consider the unique disadvantage suffered by Australia's Indigenous population is topical, given the recent High Court decision in the case of Bugmy v The Queen. In this case, the High Court recognised the circumstances of severe deprivation of Aboriginal offenders and the disproportionate incarceration of Indigenous people. The Court explicitly acknowledged that the experience of growing up in a community affected by alcohol abuse and violence is a relevant consideration in sentencing, and may reduce an offender's moral culpability for criminal behaviour. This essay suggests that mainstreaming restorative justice processes during sentencing may provide an ideal mechanism by which to bring to light circumstances of severe disadvantage and uncover underlying causes of criminal offending. In doing so, restorative justice processes will direct courts to meet the sentencing requirements of Bugmy v The Queen and will ensure that Indigenous offenders are sentenced in a manner informed by this cultural context.

13 B Steels et al 'When It’s a Question of Social Health and Well-being, the Answer is not Prison' (2009) 7 Indigenous Law Bulletin 12.
15 Kennedy v The Queen [2010] NSWCCA 260 at [53]; Fernando (1992) 76 A Crim R 58 at 62 (E); Bugmy v The Queen [2013] HCA 37 at [40].
16 Steels, above n 13, 15.
18 [2013] HCA 37.
19 Bugmy v The Queen [2013] HCA 37 at [31].
20 Bugmy v The Queen [2013] HCA 37 at [43], [44].
21 Bugmy v The Queen [2013] HCA 37.
Mainstreaming Restorative Justice in South Australia's Criminal Justice System

Inadequacy of the Current Criminal Justice System

A feature of the criminal justice system Australia-wide is the marked growth in prison populations over recent years.\(^{22}\) Such a growth in prison populations has not, however, been matched by a reduction in crime. There is little, if any, reliable evidence suggesting that increasing rates of incarceration reduce crime.\(^{23}\) Rather, studies have shown the reverse, exposing the counterproductive, harmful and brutalising nature of imprisonment.\(^{24}\) Academic commentators have observed that high rates of imprisonment may act as an indicator of the breakdown of a society’s sense of values and community.\(^{25}\) This is reflective of the correlation between the high rates of imprisonment of Indigenous people and the historic social and economic exclusion suffered by this community.

It has been suggested that the current system of criminal justice not only fails to respond to these underlying issues and work towards repairing harm, but also that it may be an aggravating factor in itself.\(^{26}\) By nature, prisons are hierarchical organisations; even the conventional architecture of prisons embodies ‘secrecy, invisibility, isolation and a lack of accountability’.\(^{27}\) These factors are likely to encourage and normalise brutality and violence amongst inmates, fail to foster a sense of personal responsibility for behaviour, and do not prepare inmates to become law-abiding members of society upon release.\(^{28}\) Not only does this system fail to address the underlying causes of criminal behaviour, it may perpetuate the cycle of offending.\(^{29}\)

The current prison system is incompatible with Indigenous culture, within which justice and accountability have their foundation on the notion of reintegrative shaming that requires the offender to face community members, family and respected elders that may play roles in the ‘sentencing’ and healing process.\(^{30}\) The one size fits all, ‘process based conveyor belt’\(^{31}\) that is the current SA prison system is likely to heighten trauma felt by Indigenous communities, rather than recognise underlying historical social exclusion and rehabilitate individual offenders. A criminal justice system augmented by the principles

---

\(^{22}\) For example, the Australian Institute of Criminology notes that the average prison population grew by five per cent annually between the years 1984 and 2003 (Dot Goulding et al ‘Restorative Prisons: Towards Radical Prison Reform’ (2008) 20 Current Issues in Criminal Justice 231).

\(^{23}\) ibid.


\(^{25}\) Coyle, above n 24, 6.


\(^{27}\) Goulding, above n 22, 232.

\(^{28}\) ibid.

\(^{29}\) ibid.

\(^{30}\) ibid.

\(^{31}\) Steels, above n 13, 15.
of restorative justice may, however, address these fundamental problems by recognising that emotional, social and health issues are intimately connected to criminal behaviour, and addressing these issues through rehabilitation.

**Part Two: The Legislative Context of Restorative Justice**

**Current Legislative Context**

This essay suggests that the issues mentioned above could be addressed through the introduction of restorative justice legislation to augment South Australia’s existing criminal justice system in which there is only limited access to restorative justice. The *Young Offenders Act 1993* (SA) provides access to ‘family conferences’, which are similar to the concept of restorative justice conferences. These, however, deal only with offences committed by juveniles. Additionally, the *Criminal Law (Sentencing) Act 1988* (SA) directs the court to consider the rehabilitation of the offender in sentencing deliberations and, in the case of Indigenous offenders, the court must be assisted by an ‘Aboriginal justice officer’ to provide advice on Aboriginal society and culture. It is also relevantly that a sentence of imprisonment may only be imposed where any other sentence would be inappropriate, having regard to the gravity or circumstances of the offence. Furthermore, in 1999 South Australia was the first state in Australia to convene an Indigenous Sentencing Court, in the form of the Nunga Court in Port Adelaide. These courts, which work to respect Indigenous culture and incorporate it into the criminal justice system, largely rely on conventional sentencing options because there is currently no legislative base for their existence. The Nunga Courts are bound, like any other sentencing court, by the *Criminal Law (Sentencing) Act* when determining sentences. Therefore, South Australia’s current criminal law legislation applies principles of restorative justice only in a limited sense.

**Proposed Legislative Changes**

Whilst some theorists, such as Braithwaite, believe restorative justice could replace the conventional criminal justice system, others acknowledge that this

32 *Young Offenders Act 1993* (SA), s10.
33 *Criminal Law (Sentencing) Act 1988* (SA), s10(1) (m).
34 *Criminal Law (Sentencing) Act 1988* (SA), s9C(5) (a).
may not be possible, and indeed should not be the ultimate aim. 37 This essay
does not recommend that the principles of restorative justice replace or usurp
the current criminal justice system in South Australia; rather, that it augments
current processes. In order to expand the application of restorative justice
and ensure that it is applied consistently, it is recommended that a new piece
of legislation be drafted with the intention of augmenting the Criminal Law
(Sentencing) Act 1988 (SA). This essay suggests that this legislation be based on
the Crimes (Restorative Justice) Act 2004 (ACT), which operates in conjunction
with the Crimes (Sentencing) Act 2005 (ACT).

The scheme under the Crimes (Restorative Justice) Act 2004 (ACT) enables
a range of criminal justice agencies and legal authorities to refer cases for
restorative justice at every stage of the criminal justice process. The legislative
scheme makes specific provision for restorative justice practices to apply
to both juvenile and adult offenders, 38 but it is being implemented in a two-
phase process. Currently, the scheme is in its ‘first phase’ and applies only to
‘less serious offences’ committed by juvenile offenders. Provisions applying to
adult offenders will come into effect as part of the ‘second phase’. 39 This essay
proposes that, given its potential benefits, the scheme be adopted in its entirety
in South Australia.

Under this scheme, restorative justice takes the form of an initial conference
between the offender and the victim or a victim’s representative, 40 whereby
parties come to an agreement that includes measures ‘intended to repair the
harm caused by the offence’. 41 These agreements can take many forms, including
apologies, community service, and programs directed to address the underlying
causes of the offending behaviour. 42 The scheme is discretionary, and allows
referring agencies to judge the suitability of the matter for restorative justice based
on the nature of each offence and the suitability of the offender to participate. 43
These agencies can choose whether to refer a matter to restorative justice, whilst
maintaining discretion to prosecute the matter within a mainstream court. 44
To allow the continuation of the court process whilst pursuing a restorative
justice solution, an offender may accept responsibility for an offence without

38 Crimes (Restorative Justice) Act 2004 (ACT), s22.
40 Crimes (Restorative Justice) Act ACT 2004, s42.
41 Crimes (Restorative Justice) Act ACT 2004, s51(1).
42 Crimes (Restorative Justice) Act ACT 2004, s51(2)(a), (b), (d).
43 Explanatory Statement, above n 6, 3.
45 Crimes (Restorative Justice) Act ACT 2004, s19(1)[b][iii].
affecting his or her capacity to plead not guilty to the offence at a later court hearing.\textsuperscript{46} In this way, both restorative justice and mainstream court processes can run concurrently.

It is recommended that South Australia adopt legislation modelled on the \textit{Crimes (Restorative Justice) Act 2004} (ACT) to operate alongside the \textit{Criminal Law (Sentencing) Act 1988} (SA). This would ensure that the principles of restorative justice are considered at all stages of the criminal justice process. Such a scheme would expand the application of restorative justice in South Australia, implement flexible approaches to the use of restorative justice, and afford the court substantial discretion apply restorative justice. Such amendments would be in line with the current legislative requirement to ‘act according to equity, good conscience and the substantial merits of the case without regard to technicalities and legal reforms.’\textsuperscript{47}

\textbf{Part Three: Potential Benefits of Restorative Justice}

\textbf{Outcomes}

Empirical evidence reveals that restorative justice is likely to have beneficial outcomes in the areas of recidivism, mental and physical health of offenders, as well as offender satisfaction. Foremost, restorative justice has been shown to reduce offender recidivism,\textsuperscript{48} with offenders who participate in restorative justice programs alongside a mainstream court process being 14 per cent less likely to reoffend.\textsuperscript{49} This is particularly evident in repeat offending for violent crime, which is 38 per cent lower in offenders who participate in restorative justice programs, compared to offenders who attend a mainstream court for the same type of offence.\textsuperscript{50} Recidivism, a term used interchangeably with reoffending or repeat offending, has been variously described as the ‘reversion of an individual to criminal behaviour’,\textsuperscript{51} the ‘return of a prisoner to custody’\textsuperscript{52}

\textsuperscript{46} \textit{Crimes (Restorative Justice) Act ACT 2004}, s20(1).
\textsuperscript{47} \textit{Criminal Law (Sentencing) Act 1988} (SA), s6(c).
\textsuperscript{50} J Hinche y ‘Restorative Justice in the Australian Capital Territory from Apprehension to Parole’ (2006) 72 Precedent 37.
Mainstreaming Restorative Justice in South Australia’s Criminal Justice System

and ‘the reappearance of a juvenile in court’. Recidivism is an issue in the SA criminal justice system, with 46.4 per cent of prison releases returning with a new correctional sanction within two years, which is slightly higher than the national average of 44.7 per cent. Overall, Australia currently sees an average of almost three arrests per criminal offender, and rates of recidivism are generally higher for Indigenous offenders. Of today’s prison population, it is estimated that 77 per cent of adult Indigenous prisoners have been imprisoned at least once before. To consider the problem prospectively, 88 per cent of male Indigenous offenders aged between 18 and 21 years are expected to be rearrested within ten years of their first arrest, compared to 61 per cent of non-Indigenous prisoners. This disparity is similar in the older population, with 48 per cent of male Indigenous offenders aged 40 years or more expected to be rearrested within ten years of their first arrest, compared to 31 per cent of their non-Indigenous counterparts. It should be noted that the differences between Indigenous and non-Indigenous offending might be the result of differential rates of detection, rather than differences in propensity for reoffending. Given these statistics, however, it is clear that an outcome of reduced recidivism is likely to be particularly beneficial to Indigenous offenders.

Principles of restorative justice place a high importance on the health and rehabilitation of offenders. This is likely to have a beneficial impact on the mental and physical health of offenders and their ability to subsequently integrate into society. Offenders who make use of treatment programs are likely to confront underlying problems and develop law-abiding coping skills, as well as enhance interpersonal skills. In particular, drug and alcohol rehabilitation may have its greatest impact on Indigenous offending, as approximately half of Indigenous prisoners are reported to link their criminal offending to drug and alcohol use. Significantly, offenders who are redirected to residential rehabilitative treatment programs, rather than being sentenced to imprisonment, are also less likely to contract hepatitis C.

Offenders experience a sense of justice and dignity when given the opportunity to be heard by a court. Research shows that court participants value highly involvement in a fair process, and they are more likely to adhere to court orders when they believe that they have been heard and respected when telling their stories. Offenders who report greater satisfaction with court process are

54 ibid 60.
55 ibid 67.
56 ibid 92.
57 National Drug and Alcohol Committee An Economic Analysis for Aboriginal and Torres Strait Islander Offenders: Prison vs Residential Treatment (Australian National Council on Drugs, August 2012).
58 ibid.
more willing to comply with the outcome of criminal proceedings, even if the outcome is adverse to them.\textsuperscript{60} Those involved in the restorative justice programs implemented under the \textit{Crimes (Restorative Justice) Act 2004} (ACT) report overwhelmingly positive outcomes from their experiences with restorative justice. Recent figures indicate that 97 per cent of all users of restorative justice programs in the ACT, including victims and offenders, report satisfaction with the services received.\textsuperscript{61} Furthermore, 99 per cent of users felt they were treated with respect, and 98 per cent felt they were able to ‘have their say’ in the process.\textsuperscript{62} These high levels of satisfaction have translated into impressive rates of compliance, with 87 per cent of offenders complying with agreements made between victim and offender at restorative justice conferences.\textsuperscript{63} These agreements include community service commitments, arrangements to undergo counselling or enter into rehabilitative programs, and financial donations made by offenders to charitable organisations.\textsuperscript{64}

### Restorative Justice and its Application to Indigenous Offenders

Restorative justice may not only provide better outcomes for individual Indigenous offenders; it may contribute to addressing the wider problem of historically entrenched social disadvantage suffered by Indigenous communities. Restorative justice is said to be concerned with the restoration of relationships between humans, but also the ‘restoration of social relationships of equality’.\textsuperscript{65} This suggests that restorative justice not only has the potential to restore relationships between victim and offender, offender and the community, but also to work towards restoring the social relationships of inequality that characterise Australia’s Indigenous community. The power of restorative justice to address wider problems of societal injustice is recognised by Braithwaite, who notes that empowering and respecting the offender through restorative justice mechanisms has the potential to restore damaged relationships within communities and prevent future injustices,\textsuperscript{66} rather than perpetuate the cycle of offending and further entrench social exclusion caused by the current criminal justice system. By recognising the inadequacy of the current criminal justice system and re-imagining our approach to criminal behaviour, progress may be made to remedy the over-representation of Indigenous people in our

\begin{footnotes}
\item[60] Ibid.
\item[62] ibid 32.
\item[63] ibid 30.
\item[64] ibid 31.
\end{footnotes}
Mainstreaming Restorative Justice in South Australia’s Criminal Justice System

criminal justice system. Doing so may identify underlying issues that cause criminal behaviour amongst Indigenous offenders. In this way, the criminal justice system has the potential to address and reverse the entrenched social disadvantage faced by the Indigenous community.

Evidence from international jurisdictions indicates that principles of restorative justice can be particularly relevant to addressing criminal behaviour within indigenous communities. The indigenous Maori population in New Zealand is disproportionately represented in the NZ criminal justice process. As with the Indigenous population of Australia, this is accounted for in part by the range of adverse social and environmental factors that put Maori people at a greater risk of repeating patterns of adult criminal activity. In 2006, the New Zealand Law Commission made recommendations to the NZ Government to expand restorative justice programs in the country, and increase court-referred restorative justice processes. The NZ experience of restorative justice has shown evidence of success, in particular with offences committed by individuals from indigenous Maori and Pacific Island cultures. It has been established that when restorative justice programs are utilised in sentencing, indigenous offenders are generally less frequently sentenced to imprisonment, imprisonment is for shorter sentences, and the rate of reoffending is reduced. This demonstrates the potential for restorative justice to address problems of historic disadvantage that are characteristic of both Maori and Aboriginal communities.

Economic Benefits

As well as offering potential benefits for individual offenders and the Indigenous community more broadly, restorative justice may present fiscal benefits for the criminal justice system. The current criminal justice system is expensive. In 2011, more than $3 billion was spent on Australian prisons, with an estimated average cost per prisoner per day in 2012–13 of $315, or $114,832 annually. Comparatively, the net financial benefit of residential rehabilitation programs compared to incarceration is $111,285 over the lifetime of an offender. Such significant savings are likely to be politically and economically attractive features of restorative justice.

---

69 ibid.
70 National Indigenous Drug and Alcohol Committee, above n 52.
Part Four: Drawbacks of Restorative Justice

Discretion versus Continuity

Whilst restorative justice offers a number of potential benefits, some features of the proposed scheme may be controversial. It is recognised that a cornerstone of our criminal justice system is consistency and predictability, and that judicial decisions should be made by reference to purely objective standards declared to the community in advance.\(^{71}\) Indeed, Braithwaite acknowledges that ‘social capital cannot flourish in a world without an infrastructure of security … that can only be guaranteed by institutions of justice’.\(^{72}\) The criminal justice system in its current form tends towards uniform consistency in sentencing. The principles of restorative justice may, however, be at odds with these objectives. In order for restorative justice to be implemented effectively, it needs to be tailored to meet the unique needs of individual offenders. The outcome of a restorative justice conference, for example, cannot be known in advance because it is often the result of an exchange between victim and offender,\(^{73}\) and the court must be allowed significant discretion in imposing sentences consistent with any restorative justice conferencing agreement if it sees fit. It can be argued that this discretionary approach to sentencing may lead to ‘unlike treatment of like cases’.\(^{74}\)

Consistency is an abstract virtue, however, that should not outweigh the primary rationales for sentencing; rather, an effective justice system should provide justice to the victim, maintain community safety and rehabilitate the offender. This is the established position at common law in Australia, held by the High Court in *Markarian v The Queen* (2005), that the correct approach to sentencing is one that weighs all of the circumstances of the offence and the offender to reach an appropriate penalty. Justice McHugh further noted that a sentence can only be the product of human judgement, and sociological variables should rightly be considered.\(^{75}\)

Genuine Participation

In reality, it may be difficult to ensure that an offender’s participation in restorative justice processes is a genuine attempt to address underlying causes of criminal behaviour. A major criticism of restorative justice is that it is a ‘soft option’ that is

---


\(^{72}\) Braithwaite, above n 1, 402.


\(^{74}\) ibid 32.

unlikely to be relevant to repeat offenders. It is not unlikely that offenders may elect to participate in restorative justice conferences with the sole motivation of achieving a reduced sentence. Whilst this is a concern, studies have indicated that restorative justice processes can be beneficial and have a meaningful impact on offenders regardless of their initial motivation for participation.

**Conclusion**

Whilst restorative justice has drawbacks, and is unlikely to deliver the ‘miraculous’ outcomes some commentators claim it will, restorative justice has the potential to provide real benefits to all those involved in the criminal justice system. Simply locating and punishing antisocial behaviour, as the current criminal justice system tends to do, does not address the root causes of criminal conduct. Statistics have proven beyond doubt that the current system is particularly detrimental to Indigenous communities and, in failing to rehabilitate offenders, is perpetuating the social exclusion suffered by Indigenous offenders. Restorative justice allows the court system to move away from restrictive notions of crime and punishment, and better connect with the communities they serve by meeting the requirement of *Bungy v The Queen* to consider the severe disadvantage suffered by Indigenous offenders.

This essay agrees that ‘there can be no justice in a world without connectedness and empathy’. It is hoped that the integration of restorative justice within South Australia’s criminal justice system may achieve a system of justice that is both empathetic to the offender and addresses the severe historical disadvantage suffered by Australia’s Indigenous community.

**Bibliography**

**Books**


---

76 Hinche, above n 50, 36.
77 Basire, above n 74, 43.
78 See, for example, Braithwaite, above n 36.
79 [2013] HCA 37.
80 Braithwaite, above n 1, 402.


**Case Law**

*Bugmy v The Queen* [2013] HCA 37

*Fernando* (1992) 76 A Crim R 58

*Kennedy v The Queen* [2010] NSWCCA 260


**Discussion Papers**


Mainstreaming Restorative Justice in South Australia’s Criminal Justice System

Government Documents


Journal Articles


**Legislation**

*Crimes (Restorative Justice) Act 2004* (ACT)

*Criminal Law (Sentencing) Act 1988* (SA)

*Explanatory Statement, Crimes (Restorative Justice) Bill 2004* (ACT)

*Young Offenders Act 1993* (SA)

**Online Resources**


**United Nations Publications**

Other


National Drug and Alcohol Committee, ‘An Economic Analysis for Aboriginal and Torres Strait Islander Offenders: Prison vs Residential Treatment’ (Australian National Council on Drugs, August 2012).
Officer-Facilitated Homicide: Assessing the Deaths of Mentally Ill Offenders as a Result of Police Practices

Benjamin A Smith

Abstract

The following essay examines three instances of mentally ill victims of homicide who were police suspects at the time of their death, and whose death was caused by the arresting attending officers. A contextual basis is provided in order to assess the conditions that facilitated those situations. This essay then addresses the questions: Are police adequately equipped to effectively manage interactions with sufferers of mental health crises? If not, why not and how might process, protocol and procedure be improved to reduce the number of deaths resulting from those inadequacies? The essay concludes that members of state police services are inadequately trained and that this inadequacy foreseeably led to the homicides. Adaptations to training modules are suggested that reflect necessary changes in order to avoid further deaths and the essay suggests the problem of mentally ill offenders be viewed as a whole-of-community issue requiring whole-of-community responses.

Introduction

Police work involves frequent interaction with mentally ill persons and a growing body of evidence suggests that these interactions are inadequately managed. The extreme consequence of this is the death of persons with mental illness (PMI); which may be avoided if police respond more effectively. This essay provides a contextual history of developments in mental health treatment and identifies some of the consequences of this. It presents three case studies and highlights common factors between them and analyses developing PMI policy. The essay advances four major arguments: 1) Police need specific, ongoing training to respond to PMI; 2) Force, particularly lethal force, should be employed only when there is a clearly defined threat to life or health, and only
as a last resort; and, 3) A whole-of-government approach should be taken to resolving mental health crises through providing the support services that police require. Supporting police involves maintaining confidence in the institution by providing for increased accountability and independent investigations.

In the 1990s Australia adopted a program of deinstitutionalisation of PMI. The rationale for this program was that those suffering had a right to be integrated with, rather than segregated from, their communities. The program was implemented in response to international principles advocating more humane treatment for PMI. Unfortunately this was not supported by an increase in community programs. A 1993 Australian Human Rights Commission (AHRC) report found services for those being treated in their communities ‘disgraceful’ and that there had been ‘virtually no retraining of psychiatric staff to work in a non-institutional setting’. As a result, many PMI did not receive adequate treatment to manage successfully in the community. This was proposed as a primary reason for the high incidence of interactions between police and PMI. The Police Foundation of Australia (PFA) noted ‘persons that are released from institutions are often released directly into the path of local law enforcement’ and that this precipitates ‘numerous violent physical confrontations between police and mentally ill persons’.

This claim is supported by a 2013 report from the Australian Institute of Criminology (AIC), which revealed that almost 50 per cent of the persons shot by police in Australia in the preceding two decades were suffering from mental health issues at the time of the incident. But when considering some of the years covered by this report individually, the number is much higher. For example, in 1994, 2001 and 2003, the figures are 67 per cent, 100 per cent and 71 per cent respectively. The report identifies these figures as conservative estimates, recognising that actual numbers may have been much higher as there were likely a number of victims with mental illness who were yet to be officially diagnosed. The report noted that the alleged offender was unarmed in 16 of these cases.

---

3 ibid.
4 Police Federation of Australia, Submission to Senate Select Committee on Mental Health, 10 May 2005 2 [3].
5 ibid., 1 [5].
6 Australian Institute of Criminology, Police Shootings of People with a Mental Illness (2013) No. 34.
7 ibid., table 1.
8 ibid., 1 [7].
9 ibid., 2 [4].
Case Studies

Individual case studies provide the opportunity to analyse how these scenarios develop. The facts of each case were collected from evidence tendered to the respective coronial inquiries.

Case Study 1: Tyler Cassidy

On 11 December 2008, 15-year-old Melburnian Tyler Cassidy experienced a crisis of mental health following a documented history of instability; this included episodes of self-harm, suicidal ideation and eventually institutionalisation. On the evening of 11 December several witnesses contacted emergency services to report interactions with a young man whose behaviour was ‘agitated and distressed’, ‘unstable’ and ‘intense’. Cassidy himself called emergency services twice throughout the evening demanding that police be dispatched. It is unclear what Cassidy’s intentions were when making these requests, but when he was encountered by members of Victoria Police (VP), he was verbally abusive and physically threatening, using knives stolen from a nearby department store. When police ordered Cassidy to drop the knives, he refused. Oleoresin capsicum (OC) spray was deployed, but proved ineffective. Cassidy advanced upon police, ignoring warning shots and repeated commands that he relinquish his weapons. The senior officer, Senior Constable Dods, was trapped and believed that he had no choice but to shoot Cassidy twice in the legs. This was followed by eight more shots, from a total of three VP members, one of which hit Cassidy in the torso and resulted in his death. From the police’s initial encounter with Cassidy to notification of his death a total of 73 seconds elapsed.

Coronial Comments on Police Response

Her Honour, Coates J, noted that in their interaction with Cassidy police failed to communicate and coordinate a planned response and that, had they done so, there was the possibility of a different result. She observed that the communications dispatch did not convey vital information to the officers. She noted that it was a failure by Dods to allow himself to be backed into a corner

---

10 Coroner’s Court of Victoria, Finding into Death with Inquest — Tyler Jordan Cassidy, 23 November 2011 15 [49].
12 ibid., 8 [4].
13 ibid., 7 [3].
14 ibid., 9 [11] [12].
15 ibid., 9 [13].
16 ibid., 9 [16].
17 ibid., 9 [18].
18 ibid., 10 [19].
19 ibid., 69 [337].
and that he should have been more aware of his surroundings, or should have been alerted to them by the other attending officers. It was further noted that police issued commands ‘in a best drum major’s voice’ whilst presuming Cassidy’s rationality, which is contrary to accepted protocols. There was also a failure to gather adequate information about the exact nature of the threat, assess Cassidy’s mental state and determine the potential for Cassidy to be reasoned with, a potential which was found to have existed. There was no use of de-escalation techniques, which is required by protocols. There were also significant issues with the investigation of the incident, including a failure to require police to give recorded evidence (thus affording them different treatment to that of the public) and a failure to perform drug and alcohol testing within required times (despite the fact that some of the officers had attended their Christmas party the night before and had been observed drinking). Additionally the investigation by VP disregarded requests from community organisations and Cassidy’s family that it be undertaken by an independent body, such as the Office of Police Integrity (OPI). Although Her Honour found that police training was inadequate, and that the four attending members had at multiple times departed from the training that had been provided, she found neither factor had substantially contributed to Cassidy’s death. It is the opinion of the author that Her Honour erred in this finding.

Case Study 2: Robert Laudisio Curti

On 18 March 2012 Brazilian national Robert Laudisio Curti suffered psychotic effects from the drug LSD and was involved in an incident in which he stole biscuits from a convenience store in Sydney’s inner city. A witness reported the incident, which was mistakenly reported by police dispatch as an armed robbery. He was also wrongly identified as attempting to steal a motor vehicle.

20 ibid., 65 [318].
21 ibid., 69 [306].
22 ibid., 66 [325] also at [321]. Her Honour notes that one of the officers was ‘the only person that evening that did not conclude that there was something significantly wrong with Tyler’.
23 ibid., 73 [361].
24 ibid., 72 [359].
25 ibid., 112 [606].
26 ibid., 120 [661].
27 Federation of Community Legal Centres Letter to the Director of the OPI, Independent Investigation into the Victoria Police Fatal Shooting of Tyler Cassidy, 12 May 2009.
29 See above n 10 [465].
30 Coroner’s Court of New South Wales, Inquiry into the Death of Robert Laudisio Curti, 14 November 2012 4 [2].
31 ibid., 5 [3].
32 ibid., 5 [4].
33 ibid., 6 [1].
34 ibid., 6 [2].
Curti was pursued initially by six,\textsuperscript{35} and later 11,\textsuperscript{36} members of NSW Police who, during the attempted arrest, deployed OC spray at least three times and tasered Curti at least 13 times.\textsuperscript{37} This, combined with the weight of ‘half a tonne’\textsuperscript{38} of officers restraining Curti (which may have led to ‘positional asphyxiation’), was declared as the likely result of his death. Although an exact cause of death could not be ascertained to a medical certainty,\textsuperscript{39} the coroner was satisfied that death would not have occurred in the absence of the activity of at least four of the officers.\textsuperscript{40} The coroner was explicit in recording that although the arrest was lawful it was unlikely that Curti presented any significant threat to the life or health of civilians or police.\textsuperscript{41}

\textbf{Coronial Comments on Police Response}

The coroner found that at least four of the police who attended the scene were reckless, careless, dangerous and excessively forceful.\textsuperscript{42} The members gave no thought to assessing Curti’s mental state.\textsuperscript{43} Other comments included an acknowledgement that Curti had committed no serious offence.\textsuperscript{44} The transcript records a separate acknowledgement that the use of ‘minimal force’\textsuperscript{45} is implicit in all police operations. Although the term now used in New South Wales is ‘reasonable force’,\textsuperscript{46} this is still framed in reference to the need to use ‘minimal force in the execution of arrests’.\textsuperscript{47} The coroner described the use of tasers and OC spray as unjustifiable and a breach of accepted procedure.\textsuperscript{48} Questions were raised about issuing junior members with tasers, the training accompanying that issue and the fact that junior members who had employed their weapons had little or no understanding of the procedures governing the use of those weapons.\textsuperscript{49} Also of associated importance is the assertion that the senior member’s failure to provide leadership and maintain objectivity was abhorrent.\textsuperscript{50}

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{35} ibid., 7 [1].
\item \textsuperscript{36} ibid., 7 [2].
\item \textsuperscript{37} ibid., 7 [2].
\item \textsuperscript{38} ibid., 23 [2].
\item \textsuperscript{39} ibid., 25–28.
\item \textsuperscript{40} ibid., 28 [2].
\item \textsuperscript{41} ibid., 21 [3].
\item \textsuperscript{42} ibid., 21 [3].
\item \textsuperscript{43} ibid., 21 [3].
\item \textsuperscript{44} ibid., 21 [3].
\item \textsuperscript{45} ibid., 18 [1].
\item \textsuperscript{46} ibid., 18 [2].
\item \textsuperscript{47} NSW Police Force Code of Practice for CRIME (Custody, Rights, Investigation, Management and Evidence) version 10 January 2012.
\item \textsuperscript{48} ibid., 20 [3].
\item \textsuperscript{49} ibid., 22 [1].
\item \textsuperscript{50} ibid., 23 [4].
\end{itemize}
\end{footnotesize}
‘contradictory’,\textsuperscript{51} ‘self-serving’,\textsuperscript{52} and ‘bearing no credibility’.\textsuperscript{53} The coroner recommended that ‘there be a review of communication procedures to ensure that signs of mental disturbance in any person the subject of a police report be communicated, and officers trained further to respond accordingly.’\textsuperscript{54}

Case Study 3: Adam Salter

Adam Salter, a 36-year-old resident of Lewisham in Sydney had suffered a nonviolent history of mental illness, leading to brief period of hospitalisation.\textsuperscript{55} On 18 November 2009 police and paramedics attended his home after Salter’s father contacted emergency services to report that Salter was engaging in self-harming behaviour by stabbing himself with a kitchen knife.\textsuperscript{56} Paramedics subdued Salter and began treating his wounds,\textsuperscript{57} which were later reported to be superficial.\textsuperscript{58} Salter’s father was taken from the room by police,\textsuperscript{59} at which point Salter freed himself from the restraint of the ambulance staff, regained the knife, (which had been removed from his person, but not the vicinity), and caused himself further harm by stabbing himself in the neck.\textsuperscript{60} The paramedics then called for police assistance, at which point Sergeant Bissett, the senior member in attendance, entered the room, yelled the words ‘taser, taser, taser’ and shot Salter in the back with her Glock firearm, causing his death.\textsuperscript{61}

Coronial Comments on Police Response

The coroner accepted the evidence tendered by paramedic staff that Salter posed no threat to anyone but himself.\textsuperscript{62} His Honour observed direct contradictions between the evidence provided by police and that provided by the paramedics, Salter’s father and the physical evidence. His Honour found that Bissett either made a tragic mistake by intending to discharge the taser but firing her handgun, or that she did intend to discharge her firearm, merely misspoke and ‘gave herself very little opportunity to assess the situation and decide whether lethal force really was necessary.’\textsuperscript{63} It was Bissett’s assertion that in fact no mistake was made. She had never intended to draw the taser

\textsuperscript{51} ibid., 15 [4].
\textsuperscript{52} ibid., 15 [2].
\textsuperscript{53} ibid., 16 [3].
\textsuperscript{54} ibid., 31 [4].
\textsuperscript{55} Local Court of NSW Coroner’s Jurisdiction, \textit{Inquest into the Death of Adam Quddus Salter}, 14 October 2011.
\textsuperscript{56} ibid., 10 [31]-[33].
\textsuperscript{57} ibid., 11 [35].
\textsuperscript{58} ibid., 4 [8].
\textsuperscript{59} ibid., 12 [40].
\textsuperscript{60} ibid., 25 [85].
\textsuperscript{61} ibid., 4 [8].
\textsuperscript{62} ibid., 24 [82].
\textsuperscript{63} ibid., 22 [75].
because she had assumed it would not work. The coroner found that given the circumstances ‘it is difficult to see that she (Bissett) gave the use of a less deadly alternative any real consideration.’ Although another officer claimed she had issued a warning, no witness (including Bissett) provided a corollary to this and, thus, the coroner rejected that assertion. It was asserted that Bissett ‘fired without much opportunity to assess what was going on and what should happen next’. His Honour found that the attending members had both provided ‘false and misleading’ evidence and were ‘almost entirely wrong’ in their description of events. The situation report was described as ‘nonsensical’ and the walkthrough was considered a ‘failure and a disgrace’. The findings of the coroner speak for themselves:

At best, the police intervention was an utter failure. Police killed the person they were supposed to be helping. They forgot to remove or secure the knife from the sink. They removed from the kitchen the very person, his father, most likely to be able to contain him. They left Adam Salter in the care of a young and inexperienced and, on the evidence of Adrian Salter and the paramedics, ineffective and unresponsive officer.

Commonalities within Case Studies

The examples detailed above are united by three major elements:

1) Police were ill-prepared to make rational and informed assessments about the situation they faced when confronted with PMI. In each of the above scenarios the inquiries noted that the police response was inadequate in that it was not the result of rational decision-making, but rather was marked by the absence of it. In the Cassidy case, the police team neglected to employ de-escalation as a technique. Further, a better assessment of the situation may have negated the need for weapons to be drawn and a better assessment of the surroundings may have prevented the need for their deployment. In the Curti case there was a clear failure to assess the true risk posed by the suspect and make an informed decision about the degree of force required. In the Salter case the coroner noted that failing to secure the offender’s weapon, failing to recognise the value of keeping Salter’s father in the room and firing without any real warning were all

64 ibid., 16 [54].
65 ibid., 24 [79].
66 ibid., 18 [61].
67 ibid., 26 [89].
68 ibid., 27 [91].
69 ibid., 27 [91].
70 ibid., 33 [111].
71 ibid., 38 [128].
failures to rationally act, rather than react. Perhaps the only means of combating the ubiquity of this factor is, as the coroner has recommended in each case, increased training and education of officers being sent into the field.

2) Police employed force without regard to proportionality, its necessity nor its probable likely result. In each of the cases it is also clear that weapons were employed disproportionately. Curti posed no risk to the public or officers; Salter posed no risk to any other person than himself, and the coroner rejected claims that he did. Cassidy presented a risk to the officers, however, it is difficult not to take seriously the coroner’s finding that perhaps even the drawing of weapons escalated the situation, after which it was difficult to resolve without an increase in aggression from both parties. What is clear from a reading of all the coroners’ reports is that necessity and proportionality in reference to the use of force was material to the death of each of the civilians.

3) The investigations, and therefore standards of accountability, were unsatisfactory in examining, assessing and regulating conduct in relation to PMI. As was noted above, there were several issues relating to the investigation of the Cassidy incident. Firstly, that the VP Homicide Squad, instead of an independent body, conducted the investigation is cause for concern. Secondly, it was noted that officers had, during the course of the investigation, covertly recorded interviews with the Cassidy family and, on this, the coroner noted that it was difficult not to draw conclusions about favourable treatment being extended to officers of VP. In the case of the Salter shooting, there was a concerted attempt by the officers involved to create a false impression of the incident in order to deflect responsibility, of which the coroner was scathing. One might further ask questions as to whether this would ever have been revealed had it not been for the presence of independent witnesses from the team of paramedics. One could further ask this question in relation to the Curti case regarding the presence of closed-circuit television footage. There is a demonstrable need for independent investigations of such incidents whenever a question arises as to the propriety of police conduct. It is not enough that independent bodies may undertake such investigations; in order to promote public confidence in the institution of policing, such bodies must undertake these investigations.

Recent Developments in Police Operations

As noted, Australia’s policy of deinstitutionalisation in the 1990s led to what was described as ‘an inordinately high number of police shootings’. This was most pronounced in Victoria (as continues to be the case), evidenced by figures
from 1990–97 which indicate that 17 of 32 fatal shootings involved PMI. It is, therefore, useful to begin a discussion of developments in police operations with the Victorian experience.

Victoria, OSTT and Project Beacon

In Victoria, police are empowered to apprehend PMI under section 10 of the Mental Health Act (1986). At present a person is detained under section 10 approximately every two hours and an incident involving the use of force occurs approximately every two and a half hours. In 1994 VP developed Project Beacon as part of the Operational Safety and Tactics Training (OSTT), which operational members are required to attend twice yearly to update academy training. Underpinning this program are the Project Beacon principles, which include ensuring:

Safety of officers, civilians, offenders and suspects is paramount.

Risk assessment is to be applied to all situations.

Violent confrontations are to be avoided wherever possible.

Use of force is to be avoided.

Minimum force is to be applied where confrontations cannot be avoided.

The safety first principle may require more resources to be be deployed, more complex planning and more time to complete.

The project and its accompanying principles were initially found to be effective in reducing both the use of force and the associated fatalities. When first implemented, Project Beacon was defined as an ‘innovative and nation leading project’ that ‘changed police practices in Victoria’. It was established in subsequent reviews, however, that there had been a shift away from the focus on reducing force and safety first. In a number of reports the OPI noted that VP demonstrated a ‘troubling cyclical pattern’ in that it seemed to implement change as a result of public controversy and, with the fading of that controversy, returned to former practices. The OPI warned VP that it was not managing the risks associated with the use of force and charged that ‘it lacked the will or capacity to implement solutions to effectively address the identified problems’. What is clear is that Project Beacon was working;

72 Office of Police Integrity, Policing People Who Appear to be Mentally Ill (2012) 17 [5].
74 Above n 10 60 [292].
77 ibid., n 69 32 [3].
78 ibid., 14.
The ANU Undergraduate Research Journal

confrontations were being more effectively and safely managed. Had this shift away from the Project Beacon principles not occurred, the Cassidy incident may have produced a different result.

The OPI made several recommendations to VP in relation to its preparedness to manage interactions with PMI. One suggestion was that it consider adopting a model based on the NSW Mental Health Intervention Team (MHIT), which it claims is the closest approximation in Australia to the successful Memphis Model, discussed below.79 One of the benefits of the Memphis Model is that it creates partnerships with mental health and emergency services. As the OPI observed in its review ‘all too often police cannot obtain the services or responses necessary to assist the individual’ and, importantly, ‘police cannot achieve better outcomes for this vulnerable section of our community by themselves. This requires a renewed commitment to a whole-of-government response to mental health.’80

Unfortunately, despite recommendations, VP declined to implement such a program.

The Memphis Model, New South Wales and the MHIT

The Memphis Model began in Tennessee in 1988, and was developed to respond to challenges presented by deinstitutionalisation in that jurisdiction.81 The Memphis Police Department joined in partnership with the Memphis Chapter of the National Alliance on Mental Illness (NAMI), mental health providers, and two local universities (the University of Memphis and the University of Tennessee) in organising, training, and implementing a specialised unit.82 In 2008, after the introduction of the Mental Health Act (2007) (which provides for police apprehensions under section 22), New South Wales adapted and introduced this program for its own use.83 The NSW MHIT includes officers with specialist training and a clinical nurse. It has resulted in the development of stronger relationships with mental health services, an increase in the use of de-escalation techniques and an increase in the confidence of MHIT-accredited officers when dealing with PMI.84 The program was reviewed by the Charles Sturt University (CSU) Graduate School of Policing and Centre for Inland Health and was noted to

79 ibid., n 69 41 [6].
80 ibid., n 69 42 [2].
82 ibid., 2.
84 Victoria Herrington, Katrina Clifford, Pota F Lawrence, Sharon Ryle, and Rod Pope Impact of the NSW Police Force Mental Health Intervention Team: Final Evaluation Report Charles Sturt University.
85 Centre for Inland Health Australian Graduate School of Policing December 2009 2–5.
have led to a marked improvement in police/PMI interactions. A major limitation, however, remains with the program’s implementation. Current figures indicate 300 operational NSW police officers have undertaken MHIT training. In a force that numbers somewhere between 15,400 and 16,200, the vast majority of NSW police are not receiving the benefit of this training. Aspirational targets of 10 per cent implementation by 2015 are insufficient. The CSU review also noted that there remain issues in relation to communicating information regarding PMI to officers in the field and recommended training for radio dispatchers. These issues are particularly important in the context of the Salter and Curti shootings. Better trained officers with a better understanding of best practice in relation to PMI and a more concerted attempt to reduce the confusion with dispatch communication may have altered the development of those incidents. Another problem lies in formulating effective inter-agency agreements. NSW Police has an inter-agency memorandum of understanding (MOU) with NSW Health and the Ambulance Service, which it updated in 2007. The CSU review was concerned that the MOU had not been further updated to reflect the legislative changes in 2007. This was preventing police from achieving its inter-agency objectives of further reducing police involvement in PMI related incidents.

**Accountability**

A number of recent changes have occurred in accountability in relation to this issue in both New South Wales and Victoria.

**Victoria**

In 2012 the OPI was replaced by the Independent Broad-based Anti-Corruption Commission (IBAC). This was hailed as an improvement in accountability and the related Act empowers IBAC to investigate either on the basis of a complaint or its own motion. It is worth recording here that the OPI also possessed these powers and determined not to use them in the Cassidy investigation. Legislative amendments are of value only to the extent that they are applied. The application of IBAC’s powers will need to be closely monitored to ensure effectiveness.

---

86 ibid.
87 NSW Police Force Operational Capacity 3 May.
88 See above n 83.
89 See above n 81 5 recommendation 1(b).
91 See above n 81 2 recommendation 3.
92 Independent Broad-based Anti-corruption Commission Act 2011 (Vic).
93 Independent Broad-based Anti-corruption Commission Act 2011 (Vic) s 64(1)(c).
94 Office of Police Integrity Act 2008 (Vic).
NSW

On 28 May 2013 and 26 June 2013 the Police Integrity Commissioner (PIC) referred officers involved in the Salter and Curti incidents to the Director of Public Prosecutions for criminal charges. Although this is a welcome sign that police in NSW are being held to higher standards of accountability, it also revealed the flawed nature of the investigative process. It is concerning that the officers involved in the Salter matter were initially recommended for a bravery award. Vigilance is required on the part of the NSW Critical Incident Investigators to ensure public confidence is maintained at all levels. This is particularly true in matters involving vulnerable members of the community.

Conclusion

This essay has outlined some issues associated with policing the mentally ill. Based on the above research it is recommended that improvements in New South Wales be built upon and delivered to all operational officers. It is further recommended that VP consider implementing a program reflecting the Memphis Model. A whole-of-government approach is required to address the complexities of policing PMI, and greater communication and cooperation between agencies is a necessary element of this. Greater emphasis must also be placed upon training and oversight of the use of force. It is unacceptable that 44 per cent of victims of police shootings are PMI; that 16 of those victims were unarmed makes it more urgent to prioritise these imperatives. Whilst there have been some welcome changes to the accountability regimes in both New South Wales and Victoria, more must be done to ensure confidence in the institution of policing and the performance of its functions.

Bibliography

Articles/Books/Reports


Australian Institute of Criminology, Police Shootings of People with a Mental Illness 34 (2013).

97 Ibid.


——, *Policing People Who Appear to be Mentally Ill* (2012).

**Legislation**

*Independent Broad-based Anti-corruption Commission Act 2011* (Vic)

*Mental Health Act 1986* (Vic)

*Mental Health Act 2007* (NSW)

*Office of Police Integrity Act 2008* (Vic)

**Other**


Federation of Community Legal Centres, Independent Investigation into the Victoria Police Fatal Shooting of Tyler Cassidy, Letter to the Director of the OPI, 12 May 2009.

*NSW Police Newsletter, ‘Mental Health Intervention Team’,* January 2008.
NSW Police Mental Health, Corporate Spokesperson Message, February 2013.

NSW Police, Operational Capacity, May 2013.


Police Federation of Australia, Submission to Senate Select Committee on Mental Health, May 2005.


Persistent Levels of Poverty in India: Urgent Action Needed to End Suffering

Harita Sridhar

This policy brief is addressed to the Government of the Republic of India and outlines and addresses causes of poverty. Poverty is highly cyclical and transferable across generations if nothing is done to prevent its spread. The national- and state-level governments have attempted to tackle poverty with many different policies and interventions, with varying levels of success. This brief identifies gross wealth and resource inequality, low literacy levels and food insecurity, and the caste system as the main causes of poverty in India and clarifies ways in which to increase the efficiency of government interventions to better address the current levels of poverty in India.

Introduction

India is the world’s most populous democracy, with a staggering population of over 1.2 billion people. Ethnic lines can be drawn down India’s 22 national languages, and also by caste. It has a history of colonisation by the Britain from the mid nineteenth century and ending approximately a century later. As of early 2012, it boasts the world’s third-largest economy, and a GDP of US$1.85 trillion.¹ The vast growth witnessed by India was triggered by the economic liberalisation reforms of 1991. Though these reforms brought India back from the brink of fiscal collapse, they also spurred the high levels of poverty and inequality existing in India today.²

Although it is estimated that there is a non-governmental organisation (NGO) in India for every 400 people,³ and despite the rapid economic growth the nation

---


has enjoyed over the last three decades, poverty has been widespread, persistent and extreme, creating issues of alarming levels of hunger,\(^4\) high rates of child labour,\(^5\) high levels of unemployment,\(^6\) and other such poverty indicators. The situation is critical, since almost one-third of India’s total population lives at or below the national poverty line.\(^7\) This equates to over 350 million individuals, in both rural and urban environments. Identifying causes of this extremely extensive social problem is challenging, as causes and consequences of poverty universally tend to be highly interlinked. This highlights the cyclical nature of poverty and how complicated it can be to break out of this loop. Some evident causes do, however, include gross wealth and resource inequality,\(^8\) low literacy levels, food insecurity,\(^9\) and the caste system.\(^10\) The extreme effects of such factors include exclusion from social, political and economic participation, starvation, low life expectancy, high infant mortality, and low literacy rates.

### Gross Wealth and Resource Inequality

India is home to more poor people than Africa, but more billionaires than England.\(^{11}\) The towering economic growth of India has vastly increased levels of inequality, and they continue to rise.\(^{12}\) This has prompted the observation that there are ‘two Indias, one shining and the other bleak’.\(^{13}\) Inequality is a cause of poverty even during periods of economic growth, as not only do the poor have the smallest share in resources, they are also excluded from benefiting from the gains of growth.\(^{14}\)

On 1 January 2013, the Indian Government instituted the Direct Benefit Transfer (DBT), which is a targeted cash transfer system designed to assist India’s poorest.\(^{15}\) Such redistributive measures are effective in reducing poverty,

\(^{6}\) ibid.
\(^{7}\) ibid.
\(^{9}\) ibid., 492–96.
\(^{13}\) ibid.
as has been seen in the successful Brazilian *Bolsa Familia* program, which is a conditional cash transfer scheme implemented to give poor families financial aid in return for keeping their children in school and acquiring basic healthcare. *Bolsa Familia* covers one-quarter of the Brazilian population, and has contributed significantly to the large drop in inequality. A major foreseeable issue for the Indian program, however, is outright corruption preventing funds from reaching the actual beneficiaries.\(^{16}\)

The state of Kerala has performed exceptionally well in the social sector. It has the highest Human Development Index (HDI) value of all the Indian states, and the lowest inequality.\(^{17}\) Its human development achievements have been credited to the active intervention of its government and good public policy.\(^{18}\) An aspect of Kerala’s success has been based on decentralising the provision of social services, which has been effective in two ways: it has enhanced service delivery by increasing transparency through local monitoring, and improved beneficiaries’ participation.\(^{19}\) When beneficiaries assist in executing social services, they have vested interests in improving the services’ efficiency and coverage. In this way, the Indian Government must work on decentralising the provision of DBT in the footsteps of Kerala to lift other states to its position.

**Low Levels of Literacy and Food Insecurity**

Low education levels are a ‘maintainer’\(^{20}\) of the poverty cycle, by keeping individuals dependent upon irregular, low-wage, high-risk employment. A lack of education also decreases people's political capacity to demand rights or have their voice heard.\(^{21}\) Much literature has been devoted to showing that education has a tremendous impact on enabling individuals to raise themselves out of poverty.\(^{22}\)

---

21 Mehta & Shah, ‘Chronic Poverty in India’, 494.
The right of children to free and compulsory education is enshrined in the Indian Constitution, but one-third of Indian children do not attend school. A national government Mid Day Meal scheme (MDM) has been operating since 1995 to provide primary school children with 100–150 grams of food grains per day. The scheme has improved enrolment, attendance and retention of children, and the likelihood of girls finishing primary school is 30 per cent higher in areas where the local school supplies a midday meal. Regarding nutrition, MDM could decrease the daily protein deficiency of primary school

---

26 Drèze & Kingdon, ‘School Participation in Rural India’, 18.
students by 100 per cent, the calorie deficiency by almost 30 per cent, and the iron deficiency by nearly ten per cent.\(^\text{27}\) Evidence suggests that a school meal improves the recipients’ intellectual abilities, which can translate into greater earning potential and therefore assist in fracturing the intergenerational transfer of poverty.\(^\text{28}\) Only 50 per cent of eligible schoolchildren receive midday meals, however, totalling about 50 million students.\(^\text{29}\) The main problems with MDM are lack of funding and monitoring provisions.\(^\text{30}\)

The Indian Government is the main body that can effectively create higher levels of literacy.\(^\text{31}\) Funding should move from higher and technical education to supporting MDM in primary schools, as higher education funding benefits the poor less than primary education funding.\(^\text{32}\) Accountability must be preserved to monitor the provision of school meals. State governments responsible for the final delivery of MDM supplies should install inspection and monitoring apparatus at all levels, and computerise records to maintain transparency.\(^\text{33}\)

Low literacy levels are also inextricably linked to perpetuating food insecurity. Food security is defined as access to enough food for an active healthy life for all people at all times.\(^\text{34}\) In contrast, food insecurity propagates the cycle of poverty, as malnutrition at childhood translates into impaired physical and mental development, lower educational attainment and higher levels of illness and disability, which eventually leads to lower adult earning.\(^\text{35}\)

The right to food is constitutional in India.\(^\text{36}\) However, India is facing alarming levels of hunger, with one in three of the world’s malnourished children living

\(^{28}\) ibid., 162.
\(^{32}\) Mukherjee, Structural Adjustment Programme, 272–73.
\(^{34}\) World Bank, ‘Poverty and Hunger: Issues and Options for Food Security in Developing Countries’, World Bank Policy Study 9275 (February 1986) 1.
in there.\textsuperscript{37} The ability of the poor to buy food is highly dependent upon food prices, which in turn are affected by the availability and production of food, and inflation.\textsuperscript{38}

The Indian Government has provided subsidised food grains for the last 70 years, with the aim of maintaining the minimum nutritional standards of the poor and shielding them from the impacts of food price increases.\textsuperscript{39} The current system — the Targeted Public Distribution System (TPDS) — provides the poorest families with up to 35 kilograms of either rice or wheat per month at heavily subsidised prices of three and two rupees respectively (approximately five and four Australian cents). Unfortunately, TPDS has a number of faults, such as targeting errors, ghost cards, unidentified households, narrow coverage and leakages/diversions of food grains.\textsuperscript{40} Only just over half the households below the poverty line are covered by TPDS, and in 2004, just six out of 14 million tonnes of food-grains released from the Central Pool (the food grain bank) were delivered to families.\textsuperscript{41}

TPDS can be improved by launching a nationwide survey to identify which households are food-insecure. The delivery system for food grains must be made more efficient and transparent by computerising all records and making them accessible to communities for monitoring.\textsuperscript{42}

The \textit{National Food Security Act}, which passed into federal law in September 2013, also advocates these measures, along with decentralisation by moving power to local institutions such as \textit{panchayats} (local governments of villages and small towns) to run the shops at which grains are sold.\textsuperscript{43}

\textbf{Caste System}

The caste system in India is a historically and structurally ingrained system of social stratification that organises individuals into hierarchical categories based on the caste of one’s family. It has been and continues to be the basis for social discrimination, though outlawed by the Indian Constitution.\textsuperscript{44}

\textsuperscript{38} Mukherjee, \textit{Structural Adjustment Programme}, 60; Ninan, ‘Economic Reforms in India’ 13; Mehta & Shah, ‘Chronic Poverty in India’, 494.
\textsuperscript{40} Program Evaluation Organisation, ‘Performance Evaluation’, ii-x.
\textsuperscript{41} ibid.
\textsuperscript{42} ibid., xi.
\textsuperscript{43} \textit{National Food Security Act} (2013) sections 18 (2)(b), 18 (2)(e).
\textsuperscript{44} Constitution of India (1950) Article 15.
categories of castes are Forward Castes (FC), Scheduled Castes (SC), Scheduled Tribes (ST) and Other Backward Castes (OBC) (the social hierarchy is ranked FC, OBC, SC then ST).  

More than four in five people in ST are multi-dimensionally poor and approximately two thirds of people in OBC and SC have low incomes. This is due to widespread and ongoing social and economic discrimination and consequent inequalities between those of higher and lower castes. For example, SC and ST people are more likely to be in bonded or casual labour, more likely to be ill and less likely to be educated. Their participation in the economic, political and social life of their community is inhibited due to prejudices in castes of higher standing. The hereditary nature of the caste system and its entrenchment in Indian social structures make it a difficult barrier to overcome. 

In 2000, only 1.4 per cent of SC and ST people aged 24–29 had a tertiary qualification, whereas 9.1 per cent of individuals in the same age bracket from FCs were so qualified. The main affirmative action taken by the Indian Government has been the reservations system, by which 27 per cent of government jobs and places in government educational institutions are reserved for those in SC, ST and OBC castes. Job reservations has increased the representation of SC/ST people in regular salaried work by five percentage points. However, drop-out and failure rates in education are higher for SC and ST students than those in other general categories, keeping poverty rates high among the lower castes. 

This happens because of the low level of basic primary and secondary education received by lower caste children. Thus, governments should focus on the early education of lower castes to create long-term sustainable growth in the number of people from lower castes in secure and high-wage employment. 

At a grassroots level, social education from an early age must be instituted in primary schools to eradicate this social impediment. For this to occur, teachers must first be educated on the backwardness of the caste system. This will lead to changes in norms and values, which could lead to a decrease in the social and

48 Mukhopadhyaya et al., Economic Growth, 118.
49 Chauhan, ‘Education and Caste in India’, 218.
51 Chauhan, ‘Education and Caste in India’, 224.
53 ibid.
economic exclusion of lower castes. As this is an ongoing cause of poverty, such social education may contribute in the long-term when addressing poverty in lower castes.\textsuperscript{54}

**Conclusion**

Poverty can be highly resistant to eradication, especially in a country as large and diverse as India. The number of people it affects makes it difficult to envisage the struggle individuals face every day. Poverty has a vast number of interlinked causes, making eradication or alleviation a complex task for governments, NGOs and other local and international bodies. Apart from the endogenous factors described in this brief, there are also a number of exogenous factors involved in poverty creation and maintenance in India. Governments, however, play the largest role in poverty reduction, as they are responsible for public policy, which extends to the largest number of people. Causes of Indian poverty, such as gross wealth and resource inequality, low literacy levels, food insecurity and the caste system must all be addressed immediately and simultaneously to instigate a comprehensive reduction in national poverty. Policies to tackle these issues must be transparent and participatory in order to engage the sectors of society who are benefiting from the public services provided. Poverty should be everyone’s problem, and economic growth is only a means to reduce hardship in countries. India must rid itself of debilitating and persistent poverty in order to decrease suffering for its people and ensure sustainable development in the twenty-first century.

Some recommendations to assist in Indian poverty alleviation:

- Move funding from higher and technical education to supporting the Mid Day Meal Scheme in primary schools.
- Install inspection and monitoring apparatus for the Mid Day Meal Scheme at all levels, and computerise records.
- Decentralise the provision of the Direct Benefits Transfer scheme.
- Improve the Targeted Public Distribution Scheme by launching a nationwide survey to identify food-insecure households.
- Make the delivery system for food grains more efficient and transparent by computerising all records and making them accessible to communities.
- Decentralise the provision of subsidised food grains by moving power to local institutions to run the shops at which grains are sold.
- Focus on the early education of lower castes.

\textsuperscript{54} Shepherd & Mehta, ‘Chronic Poverty in India’, 47.
• Institute social education from an early age in schools by firstly educating teachers.

References


Constitution of India.


Rachman, Gideon. “India is part of an upside-down world”, [http://www.ft.com/cms/s/0/8684fb7a-0630-11e2-bd29-00144feabdc0.html#axzz2Qs6I2rVuFinancial Times News](http://www.ft.com/cms/s/0/8684fb7a-0630-11e2-bd29-00144feabdc0.html#axzz2Qs6I2rVu) (24 September 2012) accessed 19 April 2013.


Gender-based Differences in Complimenting Behaviour: A Critical Literature Review

Nan Sun

Abstract

This paper provides a critical literature review of research that has examined gender differences in complimenting behaviour in English and Mandarin Chinese contexts. It first explores the functions of compliments in social life and describes the syntactic patterns and topics characterising the compliments of women and men, and then compares different types of complimenting responses used by English and Mandarin Chinese speakers. Finally, the paper discusses different interpretations of the findings offered by researchers, and identifies some research gaps for further in-depth analysis in this area.

Introduction

Compliments are common in various cultural and social contexts and are generally considered as a way to build or increase solidarity between the speaker and the addressee (Holmes, 1988). This paper analyses differences concerning compliments and compliment responses between women and men, based on previous studies conducted by researchers from two speech communities: English and Mandarin Chinese. The findings from these studies show that significant differences in complimenting behaviour across gender groups exist, including syntactic patterns and topic selections. Distributions of compliments also vary within the same gender group and across different gender groups. Women, in general, are found to give and receive compliments more often than men. Furthermore, women and men use different strategies to respond to compliments, and the gender of the interlocutor who pays the compliment is normally seen as an important factor in the English context. I consider two different explanations of the findings and suggest that, when comparing data from the two speech communities, cultural difference is a factor influencing women's and men's complimenting behaviour.
Compliment as a Speech Act and Its Functions

Holmes’s (1986b) definition of a compliment is widely used in literature:

A compliment is a speech act which explicitly or implicitly attributes credit to someone other than the speaker, usually the person addressed, for some ‘good’ (possession, characteristic, skill, etc.) which is positively valued by the speaker and the hearer.

As such, compliments can be described as ‘social lubricants’, which enable us to ‘create or maintain rapport’ (Wolfson, 1983). They are mostly used as politeness devices to express goodwill and build solidarity between the complimenter and the complimentee (Holmes, 1995), but in some contexts, as suggested by Herbert (1990), compliments serve as praise and encouragement. Herbert (1990) argues that instead of offering solidarity, compliments are likely to reflect the relationship between the complimenter and the complimentee, because praise is often directed towards the subordinate or less powerful participant.

In general, compliments are found to be multifunctional speech acts and the following summary from Holmes (1995) lists their functions as identified by different analysts:

1. to express solidarity
2. to express positive evaluation, admiration, appreciation or praise
3. to express envy or desire for hearer’s possessions
4. as verbal harassment.

These functions may, however, overlap in certain contexts. Thus, it is necessary to take other factors (such as relationships between participants and social status) into consideration when analysing the differences in complimenting behaviour between women and men.

The theoretical frameworks that have been adopted in this paper are Brown and Levinson’s politeness theory (1987) and Pomerantz’s (1978) two general conditions that govern the act of responding to compliments. Politeness theory is relevant because of its focus on redressing face-threatening acts (known as FTAs). The theory describes politeness as showing concern for two face wants: (1) negative face, which is the need not to be imposed upon; and, (2) positive face, the need to be liked or admired (Brown and Levinson, 1987). Holmes’s definition of a compliment supports the proposition that giving compliments appeals predominantly to people’s positive face, as the complimenter wants the complimentee to feel good. Under some circumstances, however, compliments can be seen as FTAs because of their evaluative function. The notion of negative
face is helpful in analysing compliment response behaviour as the complimentee is expected to choose his or her compliment response carefully, without appealing to the complimenter’s negative face. Hence the politeness theory framework is useful in evaluating compliments as politeness strategies. In analysing different strategies of compliment responses, the following two conditions proposed by Pomerantz (1978) are used:

a. agree with the speaker
b. avoid self-praise.

With these two conditions, we can understand why certain compliment responses are used to avoid FTAs in conversation. Both frameworks are valid in analysing the data collected as well as interpreting gender-based differences that are revealed in complimenting behaviour.

**Syntactic Patterns, Topics and Distributions**

One may wonder: do women and men rely on the same syntactic patterns when paying compliments? The simple answer is no. In fact, studies have found that women use more ‘personal focus’ for the act of complimenting; i.e., women tend to be involved by using the pronoun ‘I’ in their compliments (Holmes, 1988; Herbert, 1990; Sun, 2002). In other words, the syntactic pattern *I like NP* (e.g. *I like your haircut!*) is more favoured by women than men. But in Holmes’s (1988) New Zealand corpus, women are found to use the pattern *What (a) (ADJ) NP!* (e.g. *What an interesting talk!* ) significantly more often than men. In contrast, men prefer the ‘minimal pattern’ *(INT) ADJ (NP)* (e.g. *Nice outfit!* ) and use it more often than women. Similarly, in Duan and Guo’s (2009) Mainland Chinese corpus, women are also found to have a more personal focus than men when giving compliments; i.e., with ‘我(I)’ involved in their sentences. It is interesting to note, however, that few examples of third person compliments (i.e., pronouns exhibited in the compliments are not *I, we and you,* but *he, she or it;* note that a combination of a pronoun and an object is also included in this category; e.g., *Your article is really insightful!* ) exist in the English-language data collections, whereas Mandarin speakers use third person compliments (e.g., *Your suggestion is really valuable!* ) more often than English speakers. Within this category, women rely more heavily on third person compliments than men. But further study needs to be done in order to confirm this finding, as the size of the data collection from Mandarin speakers is limited, including only 64 college students from a teachers’ university in north-western China, so the sample is far from representative.
With respect to what is being complimented, previous studies have shown that
the majority of compliments focus on a few broad topics: appearance, ability, possessons, aspects of personality and performance (Holmes, 1988; Sun, 2002). Sun’s (2002) Mandarin Chinese study focuses on appearance and performance. In the New Zealand English context, Holmes (1988) reports an observably tendency for women to give and receive compliments on appearance more often than men. This finding gives credence to the stereotype that women are more concerned with appearance (Herbert, 1990; Holmes, 1995). The compliments given by men to women, in particular, have an emphasis on ability or skill, with these accounting for 44 per cent of all the compliments from men (Table 1).

Table 1: Interaction between compliment topic and sex of participants

<table>
<thead>
<tr>
<th>Topic</th>
<th>F–F</th>
<th>M–F</th>
<th>F–M</th>
<th>M–M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appearance</td>
<td>61%</td>
<td>47%</td>
<td>40%</td>
<td>36%</td>
</tr>
<tr>
<td>Ability/performance</td>
<td>20%</td>
<td>44%</td>
<td>35%</td>
<td>32%</td>
</tr>
</tbody>
</table>

Source: Amended from Holmes (1988).

Surprisingly, the findings of Sun’s (2002) study in the Mandarin Chinese context show that both women and men compliment more frequently on performance than on appearance. This contrast between the two speech communities suggests that cultural difference is another factor to consider when researchers analyse women and men’s complimenting behaviour. It must also be remembered that the participants in the Mainland Chinese studies (Sun, 2002; Duan & Guo, 2009) were college students and there are no extant studies of compliment exchanges between people of different status in the Mandarin Chinese context. In Holmes’ New Zealand data, compliments between status unequals are more likely to relate to ability/performance (54%) rather than appearance (27%); while, between status equals, compliments on appearance (57%) are much more common than on ability/performance (25%) (Holmes, 1988).

Holmes’ (1988) data show that women give (67.7%) and receive (74.3%) more compliments than men. It is worth noticing that compliments between men are relatively rare, with these accounting for only nine per cent of all compliment exchanges. In general, complimenting behaviour occurs more frequently when women are involved. Two American studies (Wolfson, 1983; Herbert, 1990) have produced similar results, as well as the two studies carried out on Mandarin Chinese (Duan & Guo, 2009; Sun, 2002).\footnote{See the section Further Discussions and Researcher’s Interpretations for further discussion of the data.}
Compliment Responses

It has been widely argued that responding to compliments appropriately is equally important in maintaining rapport or positive face of both the speaker and the addressee. In Brown and Levinson’s (1987) terms, failing to do so may ‘hurt the speaker’s face’ and hence, compliment responses can be seen as potential FTAs. Numerous studies have been done in this area, especially in the Mandarin Chinese contexts. In order to compare results from the two different speech communities more easily, I use Spencer-Oatey and Ng’s (2002) three broad categories of compliment responses, with examples in both languages.

Acceptance

Appreciation; e.g., Thanks, 谢谢 (Xièxiè)

Agreement; e.g., I like it, too, 我也挺喜欢的 (Wǒ yé tǐng xǐhuān de)

Rejection/Non-acceptance

Denial; e.g., No, I didn’t do it well, 不，做得不好 (Bù, zuò de bù hǎo)

Idiomatic rejection (an idiomatic phrase that is routinely used to refuse a compliment); e.g., 你过奖了 (Nǐ guòjiǎng le; lit. ‘you praise me too much’)

Self-praise Avoidance/ Acceptance with Amendment

Explanatory comment; e.g., A friend gave it to me, 朋友送的 (Péngyǒu sòng de)

Switch of focus; e.g., Have some more since you like it, 你喜欢就多吃一点儿 (Nǐ xǐhuān chī jiù duō chī yīdiǎnr)

The professed ideal in English is to accept a compliment graciously (Holmes, 1995) and this belief is confirmed by the quantitative data from both the New Zealand and American corpuses (Holmes, 1988; Herbert, 1990), which show that acceptance is normally considered as the correct strategy in terms of responding to a compliment. Interestingly, Holmes (1995) found that New Zealand women and men responded to compliments in similar ways. Americans, on the other hand, appeared to accept more compliments offered by men than by women, regardless of topic and status, with compliments from women being more likely to be accepted (Herbert, 1990).
Table 2: Acceptance responses

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>By sex of complimenter</td>
<td>55.3%</td>
<td>20.5%</td>
</tr>
<tr>
<td>By sex of addressee</td>
<td>28.9%</td>
<td>42.5%</td>
</tr>
</tbody>
</table>


Based on his data, Herbert (1990) suggested that the sex of the complimenter is a better indicator of the compliment response strategy used by the participants. He also pointed out that, in cross-sex interactions (i.e., F–M/M–F), women are more likely to accept compliments from men (68.6%) rather than from women (22.1%).

Many interesting findings have been reported in recent years by researchers in the Mainland Chinese community. Unlike the traditional belief that Mainland Chinese tend to reject compliments to show modesty (Chen, 1993), recent research by Cai (2012) on compliment response strategies used by Mainland Chinese college students shows that ‘acceptance strategy’ accounts for a surprising 72.3 per cent of all the compliment responses collected (from 58 males and 65 females). She attributes this change of strategy use in compliment response to the influence of what she calls ‘Western cultures’. Gender distributions in her study also show a clear tendency of women accepting more compliments than men, whereas men more often choose to opt out, either by non-acceptance or non-response. These findings are consistent with Sun’s (2002) study of college students, in which men choose to opt out almost twice as often as females.

Further Discussions and Researchers’ Interpretations

Differences have been identified in complimenting behaviour between women and men. According to Brown and Levinson (1987), paying a compliment is a prime example of a positive politeness strategy, as it appeals to people’s positive face needs; i.e., to be liked or admired by others. It follows that there exists a widespread agreement on the primary function of a compliment — to negotiate or increase solidarity (Wolfson, 1983; Holmes, 1988, 1995; Herbert, 1990). On the other hand, given that another function of a compliment is to provide positive evaluations of other people’s appearance, skill or performance, a compliment can also be seen as a potential FTA, as the complimentee may disagree on the topic of the compliment that is paid (Holmes, 1995). When it comes to interpreting the universal findings that women give and receive more compliments than men in New Zealand/American English and Mandarin Chinese, there are two popular interpretations within the current literature. One is suggested by Holmes (1988, 1995), who claims that women are primarily
focused on the positive effect of the speech act or expressing rapport and strengthening solidarity; whereas men tend to view compliments as judgments due to their evaluative function. In other words, women and men perceive the function of a compliment differently and, if men are more likely to regard compliments as FTAs, then it is natural that women have a higher probability of getting involved in complimenting exchanges. This interpretation matches the patterns shown in all the data. It should be noted that Duan and Guo (2009) also endorse this interpretation in their study on complimenting behaviour among Chinese college students.

Alternatively, in explaining the differences found in the language use of women and men, social roles connected to both sexes are relevant (Coates, 1993). Thus, some analysts, such as Wolfson (1983), have argued that the differences in complimenting behaviour across gender reflect social power, where women are usually regarded as being in subordinate positions in society. Due to the evaluative function of compliments, social approval is generally directed to the less powerful. According to Holmes ‘praise is often directed downwards from superordinate to subordinate’ (1995: 119). The tendency that is found in both languages for men to compliment women on skills and abilities can be understood as a reflection of women’s subordinate status, as women are generally seen as less powerful and influential than men. Therefore, compliments in this regard are more like encouragements rather than expressions of solidarity (Herbert, 1990; Holmes, 1995). Wolfson (1983: 91) comments that ‘the majority of compliments which occur in interaction between status unequals are given by the person in the higher position’. This assertion is confirmed by Wolfson’s data collected from American English speakers of different social statuses. The Chinese context cannot be assessed in this category as the survey of college students does not exhibit a difference in social status.

Pomerantz’s two conditions (a. agree with the speaker; and, b. avoid self-praise) are useful in analysing compliment responses and the framework has been widely used by other researchers on the same topic. According to Pomerantz (1978), complimentees are under pressure to agree with the speakers (the people who offer compliments), but on the other hand, they also want to avoid self-praise in order to show modesty. In explaining the different strategies used in compliment response by both genders, it is also essential to note that women’s language style has been claimed to be more cooperative, whereas men’s is shown to be more competitive and control-oriented (Coates, 1993). Thus, the fact that women receive and accept more compliments indicates that they are more concerned with addressing the speaker’s positive face and have a more cooperative attitude in conversations (Holmes, 1995). This is also consistent with Holmes’ interpretation that women perceive compliments as a solidarity device. By contrast, as men are more likely to see compliments as FTAs, it follows that
they tend to opt out in order to reduce the possibility of hurting the speaker’s face. Another pattern, identified by Herbert (1990) in his American corpus in which women accepted more compliments from men than from women, seems to reinforce Wolfson’s (1983) argument that women have a subordinate status in our society. Since it is generally believed that compliments usually go from higher to lower status (Herbert, 1990; Holmes, 1995; Wolfson, 1983), ‘acceptance strategy’ can be understood as the individual with less power trying to avoid disagreeing with a superior. In this case, the fact that women tend to accept more compliments from men in cross-sex interactions indicates the existence of status difference between women and men (Herbert, 1990).

Sun’s (2002) study in the Mandarin Chinese context does not reveal the gender of the addressee as being an important factor in predicting women’s compliment response strategy. His data also show that women and men use similar strategies in responding to compliments, with the majority of both women and men accepting them. Based on this, Sun (2002: 32), while acknowledging the status parity of the college students in his study, proposed that ‘women enjoy more equal rights with men in modern China’ than in the past. Sun supposes that different findings are probable given a more diverse sample size, as he believes social discrimination against women still exists globally and China is no exception; hence the traditional perception of women as the ‘low-status’ group is likely to influence people’s language use and behaviour.

In her study, Holmes (1988) pointed out that the predominance of females among the data collectors is a potential source of bias (23 of her 25 data collectors were female), and this criticism is also applied to other studies done in English contexts. She commented that patterns in complimenting behaviour may be different when data collectors include equal numbers of males and females, but she believed the imbalance would not be dramatic.

Implications

Given the considerable evidence of differences in language use between women and men in their complimenting behaviour, there is the potential for miscommunication between the sexes (Coates, 1993). Studies on complimenting behaviour across gender play an important role in helping women and men better understand each other’s communication styles, and hence reduce the likelihood of miscommunication. Thus, more studies on speech acts are strongly suggested and encouraged, especially in the Mandarin Chinese context, to fill the existing gaps in literature on gender-based differences in language use.

Cross-culturally speaking, studies on compliments and compliment responses can help second-language learners to understand the value system in the target
speech community (Wolfson, 1983). By studying how people in the target speech community perform and respond to compliments, L2 (referring to a language that is not the speaker’s native language) learners can acquire knowledge of manners and become more competent in the language. It follows that the findings of these studies also shed a light on foreign language teaching; communication breakdowns between cultures can be reduced significantly when the students are well informed of the practical knowledge or communicative strategies of specific speech acts in another speech community (Wolfson, 1983; Sun, 2002). Therefore, Wolfson (1983) suggests that future studies of speech acts must involve actual observations in realistic sociolinguistic contexts.

Conclusion

This paper has discussed the gender differences in complimenting behaviour in both English and Mandarin Chinese contexts. Overall, women are more actively involved in complimenting behaviour, both giving and receiving more compliments than men, especially on the topic of appearance. When it comes to choosing strategies to respond to compliments, women tend to accept compliments, whereas men tend to opt out more often than women. The findings from various researchers confirm the general observations and beliefs that women are more positive and supportive in conversations. But some analysts also suggest that the differences in complimenting behaviour reflect the power imbalance of gender in our society, with women as the subordinate group. English-language findings are mostly consistent with the few studies done by Mainland Chinese researchers. One surprising trend that has been identified in recent Mainland Chinese studies is that, contrary to popular belief, the most common strategy of compliment response by Mandarin speakers is acceptance. Due to the relatively few studies done in the Mandarin Chinese context, however, as well as the single methodology used by these researchers (Discourse Completion Test questionnaires), not enough data have been collected. As such, cultural difference as a potential factor in analysing compliments cannot be examined thoroughly. Further studies on the topic selections in compliments based on gender, as well as on differences in complimenting behaviour between status unequals, should be undertaken.

Bibliography


Women’s Resistance Efforts in Nazi Germany 1939–45: HerStory

Jillian Wales

Abstract

German resistance to the Third Reich garners much attention in the scholarship of Nazi Germany. Yet the resistance efforts of German women in this period have consistently been devalued, if not disregarded by historians. Examining case studies of women’s dissident behaviour in Germany between 1939 and 1945, this essay reassesses women’s resistance and argues that their efforts were significant. Despite differing motivations and backgrounds, the Rosenstrasse wives, the women of the Kreisau Circle, Maria Terwiel, Sophie Scholl, Gertrud Staewen, Erna Dubnack and others, undermined Nazi authority and overcame gendered expectations to resist in both the public and private spheres. Through protests, underground movements, aiding the enemy and non-conformity, they gave or risked their lives, resisting the Nazi regime amid an atmosphere of terror and repression. Although accounts of the crimes of Nazism, its victims, perpetrators, collaborators and passive supporters should not be downplayed, stories of women’s resistance efforts in Germany during World War II are worthy of historical recognition.

Women’s resistance efforts in Nazi Germany between 1939 and 1945 receive little scholarly attention. Indeed the majority of female resisters remain nameless, invisible in history. Accounts of the Kreisau Circle and Red Orchestra dissident groups pay fleeting attention to Countess Freya von Moltke, Countess Marion Yorck von Wartenburg and Maria ‘Mimi’ Terwiel. Herbert Baum is considered more memorable than his wife, Marianne, and Hans Scholl more impressive than his sister, Sophie. Yet these women and numerous others pitted themselves

1 The term HerStory seeks to emphasise that women’s lives, deeds and participation in human affairs have been neglected or unvalued in standard histories. See Casey Miller & Kate Swift, Words and Women (New York: Anchor Press, 1976), 146.
4 ibid.
against the totalitarian Nazi regime, deviating from expectations of femininity embodied by *Kinder, Küche, Kirche* (children, kitchen, church), in order to resist. Differing in motivations and means, these women undercut Nazi hegemony, countering the ideology and policies of National Socialism in both the public and private spheres. This essay offers an account of women’s resistance efforts in Nazi Germany by establishing a viable definition of resistance. Through the use of a number of case studies, it offers a critique of the historiography of women in the Third Reich and argues that their acts of resistance were more than a sideshow to the collapse of the Nazi regime. Against overwhelming odds and in appalling circumstances, these women stayed true to their convictions despite their small numbers.

A significant body of scholarly literature has been published regarding resistance to the Third Reich. Yet the definition, extent and success of resistance in Germany during World War II remains subject to heated debate. Matthew Stibbe in his work, *Women in the Third Reich*, interprets resistance, or *Widerstand*, to encompass actions of politically organised revolt. This narrow approach does not consider acts of dissent, non-compliance and opposition towards particular Nazi policies to constitute resistance, as they were not aimed at overthrowing the regime. In contrast to the occupied nations, such as France, there was little evidence of resistance, as defined in this narrow way, in Germany itself. As a result scholars, such as Robert Gellately and Hans Mommsen, downplay the number of German dissidents and highlight popular support for National Socialism.

---

12 As Kershaw notes, in France resistance was synonymous with attempts to oppose an invader. See for example, Large, *Contending with Hitler*, 60; Kershaw, *Nazi Dictatorship*, 184, 198.
Conversely, the broader concept of Resistenz, espoused by Martin Broszat, includes partial, passive, ambivalent and broken opposition to the Third Reich. While Resistenz undoubtedly enriches the understanding of the complex behaviour of people in Nazi Germany, encompassing their acts of conformity and non-conformity, it has been criticised for diluting the concept of resistance by regarding ‘anything short of positive enthusiasm for the regime’ as an act of resistance. Nevertheless, while Widerstand avoids the problems associated with recognising ‘almost any gesture of non-enthusiasm’ its narrowness overlooks the complex circumstances affecting German resistance during World War II.

As a consequence, this concept unjustifiably diminishes the importance of dissent. In an atmosphere of terror, fear and repression, where a trivial offence resulted not only in arrest but at times imprisonment and execution, acts of non-compliance and opposition should be considered forms of resistance. This view was espoused by former German chancellor Willy Brandt, who suggested the term ‘resister’ should be expanded to include all persons who took ‘serious risks to act illegally in a world where right ha[d] become wrong.’ As such, this research regards resistance to be positioned on a continuum between Widerstand and Resistenz, comprising behaviour ranging from private nonconformity to public acts of refusal, opposition and protest that sought to undermine the totalitarian claims of the Third Reich.

While Vera Laska vowed in 1945 to ‘bring the role of women in the cataclysm of World War Two to the attention of the public,’ historiography of World War II has largely dismissed women’s resistance efforts in Germany. Rather, this area of scholarship, for many years affected by Cold War considerations, has focused on military, conservative and political resistance networks and their male leaders. This preoccupation with men has effectively obscured and diminished

15 Richard J Evans quoted in Kershaw, Nazi Dictatorship, 205.
17 Ian Kershaw suggests that the definitional problems associated with Widerstand and Resistenz could potentially be avoided by using the term dissent as it encompasses more appropriately forms of action and the passivity of oppositional feeling. Kershaw, Nazi Dictatorship, 207.
19 Willy Brandt, quoted in Large, Contending with Hitler, 49.
the role of women, with works such as Theodore Hamerow’s *On the Road to the Wolf’s Lair*, Hans Rothfels’ *The German Opposition to Hitler* and Michael Balfour’s *Withstanding Hitler* concentrating, above all, on the resistance efforts of soldiers, bureaucrats and clergymen.23

This oversight both reflects and reinforces the widely held belief that women were passive bystanders in the male-dominated Third Reich.24 The incidence and form of women’s resistance efforts in Germany during World War II was substantially affected by deep-seated Nazi ideology grounded in patriarchy and racism.25 An ideal Nazi woman was a ‘wholesome, athletic, peasant … a domestic mother and helper to her husband.’26 As Hitler proclaimed in the 1934 Nuremberg Rally, the hereditarily healthy ‘Aryan’ woman’s world was ‘her family, her children, her house.’27 The *NS-Frauenschaft* (National Socialist Women’s Organization), the *Deutsches Frauenwerk* (German Women’s Union) as well as the *Bund Deutscher Mädel* (League of German Girls) brought large numbers of females into the Nazi fold, encouraging and elevating the duties and responsibilities of motherhood.28 Mother’s Day was commemorated as a day of national celebration, while women who bore large numbers of children were awarded the *Mutterkreuz* (Mother’s Cross).29 These constructs sought not only to enlist women’s support in preserving and expanding a ‘racially pure’ Germany, but also to further exert state control over women.30 As Jill Stephenson argues, National Socialism denied women free choice and self-determination, confining their activities and duties largely to the private sphere.31 Nevertheless, while these pervasive expectations of gender behaviour clearly limited women’s ability


28 Lillian Leigh Westerfield, *This Anguish, Like a Kind of Intimate Song*: Resistance in Women’s Literature of World War II (Kenilworth: Rodopi, 2004), 38.

29 Mothers of seven children received a gold medal, those with six, silver, and those with five, bronze. Westerfield, *This Anguish*, 38; Koonz, *Mothers in the Fatherland*, 186.

30 Westerfield, *This Anguish*, 38.

to resist the totalitarian Nazi regime, the stereotype of the passive, homebound wife arguably shielded females from official scrutiny, as authorities did not immediately suspect women of subversive activities.32

In reflecting and building upon existing studies of women’s resistance by authors such as Nathan Stoltzfus, Vera Laska and Alison Owings, this research seeks to fill this existing historiographical gap.33 Utilising case studies of women resisters, the discussion comprehensively examines the nature of women’s public and private dissident efforts in Germany from 1939 to 1945, assessing their immediate impact and overall significance. This task, however, is fraught with difficulties. Surviving records are limited, rendering it impossible to reconstruct a total picture of women’s resistance to Nazi Germany during World War II.34 Moreover, the veracity of the information contained in oral testimonies may be affected by faded memories and attempts to assuage guilt.35 Nevertheless, oral histories in conjunction with memoirs, diary entries, reports by Nazi agencies and documents produced by resisters provide extensive and valuable evidence. These sources support the conclusion that women’s resistance efforts are worthy of historical and popular recognition.

In the words of Anna Kienast, ‘the Nazis saw women as stupid and only fit to be good housewives and mothers. So my neighbour … could not imagine me to be anything else than a stupid, insignificant little woman.’36 Amid the principles of female subordination enshrined in Nazi ideology, there are numerous instances of women making a public stand against National Socialism, following their conscience in isolation and even in the face of death.37 While there existed no collective opposition, the experiences of these women demonstrate that Germany between 1939 and 1945 was not a nation of bystanders and criminals.38

A well-known act of public opposition in Germany during World War II, the Rosenstrasse Protest, illustrated that resistance was possible and could be
effective. As the inscription on the ‘Block der Frauen’ memorial to the protest reads, ‘women [stood] here, defeating death; Jewish men were free’. Until 1943 Jews in intermarriages (Aryan–Jewish marriages) were exempted from deportation. Yet, in an attempt to make Berlin Judenfrei (Free of Jews), 10,000 Jews were rounded up for deportation as part of the Fabrikaktion (Factory Action). Of them, 8000 were murdered at Auschwitz. Approximately 1700 — primarily Jewish male spouses of Aryan women — were imprisoned in a building on the Rosenstrasse to await their fate. For a week their wives protested outside, shouting adamantly, ‘Let our husbands go. We want our husbands back!’ As one witness reported, ‘the accusing, demanding cries of the women rose above the noise of the traffic.’ This protest gained the attention of the upper echelons of the Nazi regime, with Joseph Goebbels, the Reich minister of propaganda, complaining of the demonstration in his diary, ‘a large number of people gathered and in part even took sides with the Jews.’ Although the Schutzstaffel (SS) threatened to fire into the crowd and arrested ten women, over a thousand amassed in protest, arguably prompting the men’s release. While the motives behind the internment of the intermarried Jews and their later release continue to spark historical debate, it is clear the Rosenstrasse women had a considerable impact on the decisions of Nazi officials. Thus, although their opposition

---

42 Stoltzfus, ‘Dissent in Nazi Germany’, 88.
43 ibid.
44 See for example, Stibbe, Women in the Third Reich, 137; Stoltzfus, ‘Dissent in Nazi Germany’, 88; United States Holocaust Memorial Museum, ‘The Rosenstrasse Demonstration’.
45 Quoted in Stoltzfus, Resistance of the Heart, xx.
46 Quoted in Stoltzfus, ‘Dissent in Nazi Germany’, 88.
48 Estimates of the number of women present ranges up to 6000. See for example, Stoltzfus, Resistance of the Heart, 227, 238, 243; United States Holocaust Memorial Museum, ‘The Rosenstrasse Demonstration, 1943’.
49 Historian Wolf Groner contends that the Gestapo never intended to deport the intermarried Jews, but to determine their racial status and select individuals to manage Jewish institutions in place of the Volljuden (full Jews) who had been deported to Auschwitz. Yet Groner overlooks the deportation of 120 intermarried Jews to Auschwitz during the Fabrikaktion, speculates in regard to the internment of children and overemphasises adherence to Reich Security Main Office guidelines and dependent Gestapo orders. See Wolf Groner, ‘A “Historikerstreit?” A Reply to Nathan Stoltzfus’ Response’, Central European History 38(3) (2005), 460; Evan Burr Bukey, Review of Widerstand in der Rosenstrasse: Die Fabrik-Aktion und die Verfolgung der ‘Mischehen by Wolf Groner, Holocaust and Genocide Studies 21(2) (Fall 2007), 310–11; Nathan Stoltzfus, Review of Widerstand in der Rosenstrasse: Die Fabrik-Aktion und die Verfolgung der ‘Mischehen by Wolf Groner, The American History Review 112(5) (December 2007), 1629.
was personally motivated, the behaviour of the women at Rosenstrasse in publically supporting a ‘subhuman’ population at a time ‘characterized by an acceleration of violence and terror’, was significant.50 As Walter Laquer wrote, the Rosenstrasse Protest was an ‘extraordinary manifestation of courage at a time when such courage was often sadly absent.’51

The experience of Marianne Baum, a leader of the Baum Group, is justifiably lauded as a striking example of German–Jewish and Leftist resistance against National Socialism.52 Active during 1937–42, the Baum Group comprised approximately 150 individuals, many of them young communist Zionists.53 The dissident behaviour of the group concentrated on the printing and distribution of anti-Nazi leaflets, graffiti and educational evenings.54 Following the German invasion of the Soviet Union in 1941, the group’s efforts intensified, with pamphlets, slogans and the news-sheet Der Ausweg calling attention to injustice and the ramifications of ‘total war’.55 Marianne Baum’s dissident actions culminated in an arson attack on the Soviet Paradise exhibition on 18 May 1942.56 In response to the anti-Semitic and anti-Soviet nature of the exhibit, Marianne, her husband Herbert and several associates planted a firebomb which caused minor damage.57 This rare act of direct, public resistance quickly led to the discovery and demise of the group.58 On 22 May 1942 Marianne was arrested and sentenced to death.59 She was executed in Berlin-Plötzensee three months later.60 Maria Terwiel, a young half-Jewish woman associated with the aristocratic Luftwaffe officer Schulze-Boysen’s Communist group, the Red Orchestra, further exemplified Leftist resistance to National Socialism.61 As a member of the resistance group Terwiel duplicated several leaflets on her

52 The Baum Group (or Baum Gruppe) was a primarily Jewish, Leftist youth underground anti-Nazi organisation in Berlin. Cox, ‘Jewish Resistance’, 330.
54 See for example, John M Cox, *Circles of Resistance: Jewish, Leftist, and Youth Dissidence in Nazi Germany* (New York: Peter Lang Publishing, 2009), 95; Cox ‘Jewish Resistance Against Nazism’, 330.
55 See for example, Cox, *Circles of Resistance*, 107; Allan Merson, *Communist Resistance in Nazi Germany* (London: Lawrence and Wishart, 1985), 243.
58 See for example, Cox, ‘Jewish Resistance’, 330; McDonough, *Opposition and Resistance*, 8–9; Cox, *Circles of Resistance*, 133.
60 ibid.
typewriter, including a 1942 Agis pamphlet which was a call to action entitled ‘Concern about Germany’s future fills the nation.’ Maria also distributed copies of Bishop von Galen’s sermons and publications against the Soviet Paradise exhibition, which read: ‘The NAZI PARADISE. War, Hunger, Lies, Gestapo. How much longer?’ Alongside 16 other members of the Red Orchestra, Terwiel was executed on 5 August 1943. To her friend Krystyna Wituska, she was an inspiration who died ‘heroically for her ideals.’

In 1944 Hitler stated that, ‘women’s political hatred is extremely dangerous.’ In the years following the fall of Nazism and amid the Cold War atmosphere, women involved in Leftist resistance groups continued to be decried as being ‘in the service of the enemy’. Yet divesting themselves of the gendered roles imposed by the Nazi regime, these women operated in the front line of resistance, courageously pursuing their ideological commitments, ‘comforted … by the certainty that victory will be [theirs]; that [they] didn’t fight in vain for freedom.’ Although failing to overthrow the Nazi regime, the dissident actions of Marianne Baum and Terwiel put Jewish and anti-Fascist resistance on the record. Politically and ethnically motivated, the behaviour of these women demonstrated to the wider population that resistance against the National Socialist regime was possible.

Idealistic and naïve, yet fiercely committed to what they believed in, Sophie Scholl, her brother Hans and other members of the White Rose movement publicly rejected the racist militarism of National Socialism, encouraging popular resistance against the regime. From June 1942 to February 1943, the nonviolent group produced and posted leaflets in Munich, Stuttgart, Frankfurt, Linz and

---

63 See for example, Stibbe, Women in the Third Reich, 130; Noakes and Pridham, Nazism 1919–1945, 589–91.
66 Adolf Hitler quoted in Koonz, Mothers in the Fatherland, 335.
70 ibid.
Vienna. Each leaflet highlighted Nazi crimes, such as the murder of 300,000 Jews in Poland, urging fellow Germans to rise up in resistance: ‘Is your spirit already so crushed by abuse that you forget it is your right — or rather your moral duty — to eliminate this system?’ Members of the White Rose also drew graffiti on walls of public buildings, scrawling slogans such as ‘Down with Hitler!’ and ‘Freedom!’ Following the Stalingrad defeat, two more leaflets were published, denouncing the criminality of the Nazis and blaming Hitler for ‘leading German people into the abyss.’ It was when Sophie and Hans were distributing pamphlets at the University of Munich on 18 February that they were arrested. Alongside Christoph Probst, the Scholl siblings were sentenced to death by the People’s Court four days later. They were beheaded within hours.

A sister of the Scholl siblings, Elisabeth Hartnagel encapsulated best the significance of Sophie’s nonviolent resistance efforts. To her, ‘Sophie tried to show another way to the German people, that they had a choice. Sophie represents what the German people should have done.’ The life and fate of Sophie is a testament to human possibility, courage and self-sacrifice. She believed it was her moral duty to oppose tyranny and she accepted the consequences, stating during her trial, ‘so many people think like we do, only they don’t dare speak out.’ As the current German President Joachim Gauck has stated, Sophie and the White Rose ‘permit us to believe that at the time not all Germans were mute and cowardly followers’.

---

79 McDonough, *Sophie Scholl*, 158.
82 McDonough, *Sophie Scholl*, 156; Sophie Scholl, quoted in Steinhoff, Pechel and Showalter, *Voices from the Third Reich*, 357.
Historian Tim Mason argued that there was a ‘high degree of passive acceptance of the regime’ among women in the Third Reich.84 Yet within their everyday lives, numerous women developed uniquely feminine methods of opposition.85 Exploiting their roles as housewives, German women utilised the gender norms prevalent in Nazi Germany as a shield for their private dissident actions.86 Women featured prominently among those who risked their lives to rescue Jews during World War II, representing more than half of the 525 Germans given the honorific ‘Righteous among the Nations’.87 Gertrud Staewen, a member of the dissident Protestant Confessing Church, was involved in various relief efforts for Jews during the war.88 She assisted Berlin Jews in hiding: procuring passports, ration cards and accommodation for those she considered the ‘most needy among the Brothers of Christ’.89 As a ‘good Christian’ she became a ‘criminal’, feeling great pressure not because of the substantial risks associated with hiding Jews, but in knowing that she was only able to help a paltry few.90 Likewise Erna Dubnack accepted great personal risk in providing a Jewish woman food and shelter for over two years.91 Offering such assistance at the time was almost suicidal and Dubnack acknowledged that she ‘knew it was verboten’ but ‘didn’t concern [herself] with it at all’.92 Frau Naumann was her dear friend and their personal relationship rendered the danger insignificant, and the personal deprivations irrelevant.93 In the words of Leonard Gross, ‘nothing could be more miraculous than the survival of a Jew in Berlin during the last years of the Second World War.’94 With anti-Semitism considered central to the Nazi racial Weltanschauung (world view), the acts of these women in thwarting the murderous designs of the regime deserve to be recognised as significant efforts of resistance.95 As Dubnack asserted, ‘we were opponents of the Nazis’.96

86 ibid.
88 Gertrud(e) Stae wen quoted in Steinhoff, Pechel and Showalter, Voices from the Third Reich, 322.
90 From a television interview with Gertrud Staewen quoted in Renate Wind, Dietrich Bonhoeffer: A Spoke in the Wheel (Grand Rapids: Wm. B Eerdmans Publishing Co., 2002), 145; Gertrud(e) Staewen quoted in Steinhoff, Pechel and Showalter, Voices from the Third Reich, 323.
91 Owings, Frauen, 432.
92 Erna Dubnack quoted in Owings, Frauen, 438.
93 Owings, Frauen, 432.
96 Erna Dubnack quoted in Owings, Frauen, 444.
As members of the so-called Kreisau Circle, the actions of Freya von Moltke and Marion Yorck von Waternburg should not be dismissed.97 Meeting at the Wartenburg apartment in Berlin, or at the Moltke estate, the Kreisau group served as a ‘spiritual counterpoint to the war’ from 1940 to 1944.98 United in an attempt to plan for a future Germany, the group discussed how to establish a just and humane political and social order after the collapse of the regime.99 The substance of their debates came under the banner of opposition to National Socialism.100 Although criticised as too hypothetical and philosophical to constitute resistance, the actions of the group in preparing for the fall of Nazism were tantamount to treason.101 Von Moltke knew the likely punishment was death, stating that ‘we lived with the feeling that what we did was right. That was much stronger … we weren’t afraid.’102 Following a failed attempt to assassinate Hitler on 20 July 1944, the Kreisau Circle was exposed and, while the extent of the group’s involvement in the attempt was unclear, many members were executed, including its founders Helmuth James von Moltke and Peter Yorck von Wartenburg.103 Although less prominent than their husbands, Freya and Marion belonged to the core of the group.104 Freya concealed the circle’s papers, while Marion delivered messages.105 Both were intimately involved with the group’s discussions and actions, with Marion spending three months in prison following the July plot.106 At a Berlin church in 2004, Freya said of the Kreisau group, ‘even though we had no success … and even though we were weak, we kept European humanity alive in Germany’.107 Although the resistance efforts of the Kreisau wives were primarily in aiding their husbands, their contributions were undoubtedly significant.108 As Marion noted, ‘I have

97 The Kreisau Circle was the name given to this group of German resisters by the Gestapo. Van der Vat, ‘Countess Freya von Moltke Obituary’; Muriel Cormican, Review of The Power of Solitude: My Life in the German Resistance by Marion Yorck von Wartenburg, South Atlantic Review 67(3) (Summer 2002), 108; Freya von Moltke, Memories of Kreisau and the German Resistance, tr. Julie M Winter (Lincoln: University of Nebraska Press, 1997), xi.

98 Freya von Moltke quoted in Owings, Frauen, 253.


100 Yorck Von Waternburg, The Power of Solitude, 34.

101 Owings, Frauen, 253.

102 Freya von Moltke quoted in Owings, Frauen, 257.


105 See for example, Yorck Von Waternburg, The Power of Solitude, 37; van der Vat, ‘Countess Freya von Moltke Obituary’.

106 Marion Yorck von Wartenburg quoted in Steinhoff, Pechel and Showalter, Voices from the Third Reich, 398.


108 Von Moltke, Memories of Kreisau and the German Resistance, xi.
in fact the sense that Peter and Helmuth and Adam Trott and particularly Hans Haeften and all the others, could not have done everything they did without their wives.'

Even though women who engaged in non-conformist behaviour cannot be considered resisters in the strict sense, they nevertheless showed contempt for the National Socialist regime and its policies at a time when it was dangerous to do so. Frau Margarete Fischer and Frau Maria von Lingen both listened to foreign broadcasts during the war, even though they were aware that it was ‘forbidden’, with Maria interrogated three times by the Gestapo. The experiences of Ilse Demme and Marianne Elise K highlight the risk of committing these minor offences. A first degree Jewish Mischling, Demme was sentenced to three years’ imprisonment for alleged violations of the Heimtückegesetz (Malicious Practices Act) on 18 August 1942. It was claimed she had listened to ‘enemy radio programs’ and passed on information which ‘undermin[ed] the people’s trust in the political leadership … through maliciously spiteful comments.’ Marianne Elise K was similarly convicted of making ‘spiteful remarks’ and executed in 1944 for telling this joke: ‘Hitler and Göring are standing on top of Berlin’s radio tower. Hitler says he wants to do something to cheer up the people of Berlin. “Why don’t you just jump?” suggests Göring.’ Collectively, these acts of women’s non-conformity in Germany during World War II undercut Nazi hegemony in a small but noteworthy way. In a society where ‘even just a dumb remark’ could get you shot, refusing to give the ‘Heil Hitler’ salute, listening to Allied radio frequencies and making anti-state remarks arguably constituted significant psychological resistance to the ‘ideology and policies of National Socialism’.

110 See for example, Stibbe, Women in the Third Reich, 128; Kitchen, Nazi Germany at War, 237; Kershaw Popular Opinion and Political Dissent, 183–84; Kweet, ‘Problems of Jewish Resistance Historiography’, 41; Jill Stephenson quoted in Stibbe, Women in the Third Reich, 132; Jacobsen, Germans against Hitler, 162.
111 Margarete Fischer and Maria von Lingen quoted in Owings, Frauen, 5, 122.
113 ibid.
115 Large, Contending with Hitler, 51.
As these case studies demonstrate, the resistance efforts of women throughout Germany between 1939 and 1945 should no longer be overlooked. Whether they were aristocrats, communists, Christians, Jews or students, they overcame barriers specific to their gendered role in society and resisted the Nazi regime. Motivated by personal, ethnic, ethical and political reasons, these women courageously opposed National Socialism, protesting and encouraging popular opposition in the public sphere as well as rescuing Jews, conspiring against the government and resisting the penetration of Nazism through non-conformity in the private sphere.\footnote{117} Although their efforts failed to overthrow the regime and restore justice, these women played an invaluable role in the secret war against the Nazis.\footnote{118} In giving or risking their lives in opposition to the Third Reich during World War II, these heroines deserve historical recognition for their efforts in their own right. As the ‘Memorial to the German Resistance’ reads:

You did not bear the shame. You resisted. You bestowed an eternally vigilant symbol of change, by sacrificing your impassioned lives for freedom, justice and honour.\footnote{119}

**Bibliography**

**Primary Sources**


\footnote{117}{Koonz, *Mothers in the Fatherland*, 342; Kweit, ‘Problems of Jewish Resistance Historiography’, 41; See for example, Cox, *Circles of Resistance*, 5.}


**Modern Editions**


Newspapers


Secondary Sources

Books


Kaplan, Marion A. *Between Dignity and Despair: Jewish Life in Nazi Germany*. Oxford University Press, 1998.


**Articles**


Websites


Other


Comparing the Energy Consumption of Logistics Chains in Traditional Retailing and E-commerce for Popular Products in the Context of China

Wenjia Wang

Abstract

E-commerce has transformed the way in which people shop; and the introduction of modified logistics chains has profoundly influenced the retail industry. Changes in the logistic chains of e-commerce created uncertainty regarding the energy efficiency of this new business model. This paper compares the energy efficiency of logistics chains in China between traditional retailing and e-commerce models across a range of products that are mostly bought online. The results show that the percentage of airfreight used for transportation and the distribution distance in the e-commerce model has a strong influence on the overall energy efficiency of e-commerce logistics chains.

Introduction

In the past decade, information and communication technology (ICT), including the rise of online shopping, has fundamentally transformed people’s lives. The convenience of online shopping and the variety of products available on the Internet have converted many store shoppers to home shoppers. According to a survey conducted by the Australian Communications and Media Authority (ACMA) regarding Australians’ participation in e-commerce activities, 88 per cent of respondents performed at least one e-commerce activity in the past six months (ACMA 2010). In China, the number of online shoppers has reached 220 million (Redant n.d.).

With the proliferation of e-commerce, however, there is increased uncertainty regarding energy efficiency — in terms of the energy consumption incurred in the logistics chain for a purchase — in relation to the new business model. E-commerce is expected to influence a range of man-made systems of industrial
infrastructure, including systems of manufacturing, transportation, packaging and warehousing (Abukhader 2003). For the retailing industry, the influence of ICT is profound. New logistics chains have been introduced into a model of e-commerce in which product and information flows will be significantly different from traditional retailing (Berger 2003).

This paper investigates the products most frequently purchased on the Internet and compares logistics chain energy consumption in traditional retailing and e-commerce models for retailing these products. Ideally, the methodology proposed in this paper will serve as a useful tool for the generalised study of the impacts of e-commerce on the retailing industry in terms of logistics chain energy consumption.

**Comparison Framework**

**Classification of Product Types**

According to the China Internet Network Information Centre (CNNIC) the most frequently purchased online merchandise in China is (in order of purchase frequency) (CNNIC 2013):

- clothing/shoes
- general merchandise
- computer equipment, communication devices and other digital products
- white goods
- books and audio-visual products
- health and beauty items
- groceries

Clearly these products vary from each other in terms of average size, weight and monetary value. Each of these features has an impact on the structure and energy consumption of the logistics chain during retailing. Thus the products are further categorised into 6 groups based on average size, weight and monetary value, as shown in Table 1.
Comparing the Energy Consumption of Logistics Chains in Traditional Retailing and E-commerce

Table 1: Classification of Products Based on Product Type and Characteristics

<table>
<thead>
<tr>
<th>Items</th>
<th>Size</th>
<th>Weight</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clothing/shoes</td>
<td>Small</td>
<td>Light</td>
<td>Low</td>
</tr>
<tr>
<td>Audio-visual products</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health &amp; beauty items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General merchandise</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Groceries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer equipment</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Communication devices and other digital</td>
<td>Small</td>
<td>Light</td>
<td>Medium</td>
</tr>
<tr>
<td>products</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White goods</td>
<td>Large</td>
<td>Heavy</td>
<td>High</td>
</tr>
<tr>
<td>Books</td>
<td>Small</td>
<td>Medium</td>
<td>Low</td>
</tr>
</tbody>
</table>

Source: Author’s analysis.

The determination of the average size and weight ranges of the products is based on packaging advice provided by DHL Size and Price Guide (DHL n.d.). Accordingly, the product can be classified as small if it can fit into a 337 x 322 x 180 millimetre (mm) box, medium if it can fit into a 481 x 404 x 389 mm box and large if it cannot fit into a 481 x 404 x 389 mm box. The product could be classified as light if it weighs less than five kilograms (kg), medium-weight if it weighs between five and 20 kg and heavy if it weighs more than 20 kg.

As for the numerical range of monetary values of the products, after an investigation of the price ranges for the listed products, a price range of ¥100–500 (~AUD 17.5–88) for low-value products, ¥500–2000 (~AUD 88–500) for medium-value products, and ¥2000 (~AUD 500) and higher has been assumed for high-value products.

For simplicity of calculation, the upper weight limits and the mean values of the price range were used. For heavy and expensive products, i.e. white goods, an average weight of 40 kg and a value of ¥6000 (~AUD 1050) was assumed, based on available data (City of York Council 2013).

Three Logistics Chain Models

A logistics life cycle assessment (LCA) is conducted for each type of product to streamline the components of the logistics chains in both traditional retailing and e-commerce models. Based on the LCA, three logistics chain models are proposed for the listed products.
Logistics Chain Model for Regular Products (the Generic Logistics Chain)

All the listed products, except white goods and groceries, fit into the logistics chain model as depicted in Figure 1. The products in the two retailing models begin at the manufacturer, from where they are later shipped to the wholesaler’s warehouse. After the wholesale warehousing stage, the logistics chain in the two retailing models begin to separate. In the traditional retailing model, on receiving the order from the retailer, the products are transported to either the retailer’s warehouse or directly to the retail store, and the customer must go to the retail store to pick up the product. In the e-commerce model, the customer places an online order at home. The online retailer receives the order through the data centre of the e-commerce platform and the product is dispatched from the wholesale warehouse to the final distribution centre. Here the product is collected and sorted for last-mile delivery, with a possible stop at the online retailer’s secondary warehousing facility. The e-commerce logistics chain concludes with last-mile delivery to the end customer.

Figure 1: Logistics Chain Model for Large-Sized Products (White Goods)

Source: Author’s own work.
Figure 2 depicts the logistics chain for large-sized products, such as white goods. The logistics link from the manufacturer to the wholesale warehouse is similar to that for regular products. For both traditional retailing and e-commerce models, however, a secondary warehousing stage at the retailers is necessary, since the lead time taken for the large-sized products to be delivered to the retailer is usually longer. The product flow separates after the secondary warehousing stage. In the traditional retailing model, the products are delivered to the retail store and the customer travels to the store to purchase the product. The final delivery of the product to the customer is made by the retailer. In the e-commerce model, the customer places the order online. On the order’s arrival, the online retailer dispatches the product from its warehouse. After going through several distribution centres, the product reaches the final local distribution centre and the last-mile delivery of the product is made to the customer.

Figure 2: Logistics Chain Model for Grocery Products

Source: Author’s own work.

Figure 3 depicts the logistics chain for grocery products. As grocery products are often perishable, it is not realistic for a customer to order the products from a non-local e-commerce company. Most of online grocery retailers are multi-channel players; that is, they own retail stores and at the same time they are able to take online orders and deliver the groceries to their customers. The only difference between traditional retailing and e-commerce models for grocery products is in the final logistical link, which is fulfilled by customer travel in the traditional retailing model, versus a last-mile delivery by the retailer in the e-commerce model.
Figure 3: Grocery Retailing Logistics Chain in Online Retailing Model

Source: Author’s own work.
The three logistics models proposed above illustrate that the initial stages of the logistics chains for both retailing models are very similar. Thus, the energy consumption at these initial stages is assumed to be similar and can be excluded from the comparative study. It is hard to draw a comparison between the energy consumption of the two secondary warehousing stages due to insufficient information; hence, we assume that energy consumption at the retailer’s warehousing facilities is the same for both retailing models and it can be excluded from the comparative study. In the e-commerce model, we also assume that the energy spent on the product in the distribution centre is small and can be ignored for the comparative study, considering the massive amount of goods that the distribution centre needs to process on a daily basis.

For a generic logistics model, differences in energy consumption between transport links are from the distribution stage, customer travel and last-mile delivery. The non-transportation energy consumption differences are from home and data centre energy usage, and additional packaging materials in the e-commerce model and the retail store in the traditional retailing model. For large-sized products, last-mile delivery is excluded from the scope of this study as it is made in both retailing models. As the packaging methods are similar in both retailing models for large-sized products, the usage difference is assumed to be small and can be disregarded. The grocery logistics model assumes that the retailers in both models are multi-players and, therefore, the retail store and distribution stages are the same and excluded from the study.

In summary, the logistics links that are considered in this project for the three logistics models are summarised in Table 2.
<table>
<thead>
<tr>
<th></th>
<th>TRADITIONAL RETAILING</th>
<th>E-COMMERCE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Retail store</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Distribution</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Additional packaging</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Data centre</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Energy consumption at home for placing orders online</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Customer transport</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Last-mile delivery</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Product Type</strong></th>
<th>Regular products</th>
<th>Large-sized products</th>
<th>Groceries</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Regular products</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Large-sized products</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Groceries</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Source: Author’s analysis.
Comparing the Energy Consumption of Logistics Chains in Traditional Retailing and E-commerce

Methodology

All energy consumption was calculated and presented in megajoules (MJ). Unit abbreviations and the conversion between different energy units are presented in appendices 1 and 2.

Retail Store Energy Use

Data on energy consumption of retail stores was obtained from a survey done by China Chain Store & Franchise Association (CRA 2011). In the survey, it is suggested that the energy consumption of a specialty store is about 70 kilowatt hours (kWh) to 150 kWh for every ¥ 10,000 (~AUD 1755) of sales generated. The monetary value of the product can be assumed to reflect the energy usage at the retail store. A mean value 110 kWh/¥ 10,000 (~AUD 1755) is used for calculation. The expression for retail store energy use is shown as (the result is converted into MJ):

\[ E_{\text{retail store}} = \frac{P_{\text{unit}}}{¥ 10,000} \times 110 \text{ kWh} \times 3.6 \text{ MJ/kWh} \]

Energy Consumption at Home for Placing Orders Online

The energy consumption at home for placing orders online comes from four major sectors: computer, lighting, heating/cooling and the data transfer on completing an online transaction (Weber et al. 2008), as shown in the expression:

\[ E_{\text{home}} = (P_{\text{computer}} + P_{\text{lighting}} + P_{\text{heating/cooling}}) \times t_{\text{browsing}} + E_{\text{data transfer}} \]

Data on lighting is taken from the Standard for Lighting Design of Buildings specified by the Ministry of Housing and Urban-Rural Development of China (MOHURD), which suggests that the standard domestic lighting power be seven watts per square metre (W/m²) (MOHURD 2004). Data on energy use of home climate control systems is taken from China Building Energy Efficiency Development Report, which suggests that the air-conditioning energy consumption for household climate control in 2004 was 2.7 kWh/m² (Tsinghua University 2009). To convert the annual energy use into power, we assumed a usage pattern for air-conditioners of ten hours per day, 365 days per year. The power usage of the climate control system is calculated to be 0.74 W/m² (2.7 kWh/m² divided by 3650 hours of annual operating time). From the same report, it is suggested that the average household area per capita in urbanised regions is 17.7 m² in China, so the power usage of lighting and climate control systems at home are calculated to be 123.9 W (7 W/m² × 17.7 m²) and 13.1 W (0.74 W/m² × 17.7 m²).
To calculate computer energy use, energy consumption in both the operation phase and the production phase needs to be considered. Computers are different from other household appliances as 80 per cent of their life cycle energy is consumed during the production phase, while only 20 per cent accounts for the energy consumed during the operation phase (Consciousness 1993). We assumed that the power rating of an average PC is 100 W. As for energy use in the production phase, a study done by Williams (2004) indicated that the production energy of a certain desktop computer is 6400 MJ. Assuming the usage pattern for a computer is three hours per day with no standby, 365 days per year, for three years, the equivalent production power of the computer is calculated to be 542 W (6400 MJ divided by the total operating time). Including operating power (100 W), in total the power rating of a computer is 642 W. The average data usage incurred for completing one online transaction is assumed to be 1 Mb (megabyte). A study done by Taylor and Koomey (2008) suggested that the energy consumption for Internet traffic is \( \sim 9\text{–}16 \text{kWh/GB} \) of data. Using the mean value (12.5 kWh/GB), the network energy consumption for an online transaction with 1 Mb of data transfer is calculated to be 0.045 MJ.

The time spent on placing an online order is closely related to the monetary value of the product. It is suggested that customers dwell on purchasing decisions for longer if the products are in the high price range (Greenleaf & Lehmann 1995). For this study the time spent placing an online order to buy products with low, medium and high value is assumed to be 30, 45 and 60 minutes respectively.

**Data Centre Energy Usage**

To calculate energy use at the data centre, firstly the power usage effectiveness (PUE) factor is introduced as a measure of how efficiently the ICT facilities in the data centre extract energy from the total energy input of the data centre.

\[
PUE = \frac{E_{\text{total data centre}}}{E_{\text{ICT facilities}}}
\]

Accordingly, if PUE and the energy usage of ICT facilities are known, the total energy used in the data centre can be calculated. Dividing the total energy usage by the total number of transactions within the same time frame, the data centre energy usage for completing one transaction can be calculated, as shown in expression:

\[
E_{\text{data center}} = \frac{1}{N_{\text{total tranctions}}} \times PUE \times E_{\text{ICT facilities}}
\]

Energy consumption at the data centre is independent of product types. Online retailers are assumed to be operating on Taobao.com, which is the biggest e-commerce platform in China. A general PUE value of 2.5 in China and a total annual energy consumption of the ICT facilities at the data centre of Taobao.com
Comparing the Energy Consumption of Logistics Chains in Traditional Retailing and E-commerce

of 78,840 megawatt hours (MWh) was used (Wang & Liu 2012). The number of daily transactions made on Taobao.com is 8 million (iResearch 2012). As a result, the energy usage for processing one transaction at the data centre is 0.2 MJ.

Distribution

The distribution stage refers in the traditional retailing model to the transportation of the product from the wholesaler’s warehouse to the retail store, and in the e-commerce model, to the final local distribution centre. With no better information available, the distribution distances in the two retailing modes are assumed to be similar. For calculation purposes, the distribution distances are estimated to be 1000 kilometres (km), which is equal to the distance between Beijing and Shanghai.

The transportation method during distribution is assumed to be determined by the size and weight of the product. For small- to medium-sized and light- to medium-weight products, both ground delivery and airfreight are applicable. For large-sized and heavy products, however, ground can be the only option as the product might exceed the size limit of air transportation, or this option is not cost-effective. Transportation in the traditional retailing model is assumed to have been completed with 100 per cent ground delivery via medium truck. For the e-commerce model, small- and medium-sized products are assumed to be transported between the local logistics facilities (from the wholesale warehouse to the despatch air hub, from the arrival air hub to final distribution centre) via light truck for 200 km, and between the regional logistics facilities (from the despatch air hub to the arrival air hub) via airfreight for 800 km. For large-sized products, e.g., white goods, the distribution link is assumed to be completed with 100 per cent of ground delivery via medium truck in the e-commerce model. The energy intensities of the transportation methods mentioned above are listed in Table 3.

Table 3: Energy Intensity and Load Capacity of Transportation Methods

<table>
<thead>
<tr>
<th>Transportation method</th>
<th>Load capacity</th>
<th>Energy intensity (e)</th>
<th>Fuel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light truck</td>
<td>5 tonnes</td>
<td>24.2 litres (L) of fuel/100 km</td>
<td>Diesel</td>
</tr>
<tr>
<td>Medium truck</td>
<td>10 tonnes</td>
<td>27.6 L of fuel/100 km</td>
<td>Diesel</td>
</tr>
<tr>
<td>Air carrier</td>
<td>N/A</td>
<td>22.6 MJ/t·km</td>
<td>Jet Fuel</td>
</tr>
</tbody>
</table>

Source: Jia et al. 2010; Wu 2007.

To calculate the energy consumption dedicated to transporting one unit of a product, the weight ratio of one unit of the product to the truckload is required.
For ground delivery, the truck loading factor (LF) was taken into consideration, which is defined by the ratio of the weight (or volume) carried by the vehicle times the distance travelled by the weight (or volume) capacity of the vehicle times the distance travelled (Williams 2002). In China, the loading factor of trucks is around 60 per cent (Pan n.d.). Thus the energy use for transporting one unit of a product can be calculated using the following expressions, where the density of diesel ($\rho$) is 0.85 kg/L and the high heat value (HHV) of diesel is 42652 kJ/kg:

$$E_{\text{ground}} = \frac{m_{\text{unit}}}{m_{\text{load capacity}}} \times LF \times S_{\text{distribution distance}} \times e_{\text{truck}} \times \rho \times \text{HHV}_{\text{diesel}}$$

$$E_{\text{air}} = m_{\text{unit}} \times S_{\text{distribution distance}} \times e_{\text{air}}$$

### Customer transport

In the traditional retailing model, the distance customers are willing to travel is dependent on the type of product they are going to purchase, as shown in Table 4. Customers also choose the type of transportation based on round-trip shopping distance. Table 5 shows the percentage of customers choosing different types of transportation for different distances. Notice that the sum of percentage under each transportation method for a travel distance does not necessarily result in 100 per cent, as the customers might use multiple transportation methods for a single trip. Also, the customer might purchase more than one product in a single trip. The percentage data in Table 5 can be translated into the percentage of the travel distance covered by a certain transportation method. The energy intensity for each transportation method is summarised in Table 6. To calculate the customer transport energy usage dedicated to buying one unit of a product, a multi-purpose factor (MF) of 0.5 for a shopping trip is assumed, indicating that 50 per cent of the shopping trip energy is used for buying one unit of the product. Frequently purchased products are classified in terms of product types in Table 7. The expression for calculating the energy consumption at the customer transportation stage is shown below, where $f$ stands for the percentage of the shopping trip covered by a certain transportation method, and $e$ stands for the energy intensity of that transportation method:

$$E_{\text{customer transport}} = \sum (f \times e) \times S_{\text{shopping distance}} \times MF$$
Table 4: Customer Transport Distance for Each Type of Product

<table>
<thead>
<tr>
<th>Product type</th>
<th>Customer travel distance</th>
<th>Round trip distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture &amp; electronics</td>
<td>5.3 km</td>
<td>10.6 km</td>
</tr>
<tr>
<td>Necessities</td>
<td>1.9 km</td>
<td>3.8 km</td>
</tr>
<tr>
<td>Garments</td>
<td>6.5 km</td>
<td>13 km</td>
</tr>
<tr>
<td>Groceries</td>
<td>0.9 km</td>
<td>1.8 km</td>
</tr>
</tbody>
</table>

Source: Chai, Weng & Shen 2008.

Table 5: Percentage of Customers (f) using certain Transportation Methods to the Retail Store — Based on Different Types of Products (Can Also be Interpreted as the Percentage of the Travel Distance (f) Covered by a certain Transportation Method)

<table>
<thead>
<tr>
<th>Round trip distance</th>
<th>f (Walking)/%</th>
<th>f (Cycling)/%</th>
<th>f (Bus)/%</th>
<th>f (Subway)/%</th>
<th>f (Car)/%</th>
<th>f (Others)/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.6 km</td>
<td>0</td>
<td>5.10</td>
<td>30.37</td>
<td>12.21</td>
<td>61.26</td>
<td>0.86</td>
</tr>
<tr>
<td>3.8 km</td>
<td>4.50</td>
<td>30.04</td>
<td>30.51</td>
<td>0.81</td>
<td>30.17</td>
<td>2.33</td>
</tr>
<tr>
<td>13 km</td>
<td>0</td>
<td>5.10</td>
<td>30.37</td>
<td>12.21</td>
<td>61.26</td>
<td>0.86</td>
</tr>
<tr>
<td>1.8 km</td>
<td>56.41</td>
<td>29.08</td>
<td>3.05</td>
<td>0.02</td>
<td>7.60</td>
<td>3.54</td>
</tr>
</tbody>
</table>

Source: Huang, Song & Tao 2008.

Table 6: Energy Intensity of Customer Transportation Methods

<table>
<thead>
<tr>
<th>Transportation method</th>
<th>Energy intensity (e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>4.0 MJ/vehicle* km</td>
</tr>
<tr>
<td>Bus</td>
<td>651 kJ/person* km</td>
</tr>
<tr>
<td>Subway</td>
<td>117 kJ/person* km</td>
</tr>
<tr>
<td>Walking/cycling</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7: Summary of Product Types

<table>
<thead>
<tr>
<th>Product type</th>
<th>Items</th>
<th>Round trip distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Furniture &amp; electronics</td>
<td>White goods, Computer equipment, Communication devices, Other digital products, Audio-visual products</td>
<td>10.6 km</td>
</tr>
<tr>
<td>Necessities</td>
<td>Books, Health and beauty items, General merchandise</td>
<td>3.8 km</td>
</tr>
<tr>
<td>Garments</td>
<td>Clothing/shoes</td>
<td>13 km</td>
</tr>
<tr>
<td>Groceries</td>
<td>Groceries</td>
<td>1.8 km</td>
</tr>
</tbody>
</table>

Source: Author’s analysis.

Last-Mile Delivery

For the last leg of the logistics chain in the e-commerce model, the delivery distance is assumed to be 100 km between the final distribution centre and the customer and the transportation to be via light van, with the energy intensity shown in Table 8. Similar to the calculation of distribution energy use, the weight of the product influences the energy consumption at this stage. A loading factor (LP) of 60 per cent was applied. With the representative weight specified for light, medium and heavy products, the last-mile delivery energy consumption can be calculated using the expression:

\[ E_{\text{last mile}} = \frac{m_{\text{item}}}{m_{\text{load capacity}}} \times LP \times S_{\text{delivery distance}} \times e_{\text{transportation}} \]

Table 8: Energy Intensities of Last-Mile Delivery Transportation Methods

<table>
<thead>
<tr>
<th>Transportation method</th>
<th>Load capacity</th>
<th>Energy intensity</th>
<th>Fuel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light van</td>
<td>1 ton</td>
<td>12.5 L of fuel/100 km</td>
<td>Gasoline ((=0.72) kg/L HHV = 44000 kJ/kg)</td>
</tr>
</tbody>
</table>

Comparing the Energy Consumption of Logistics Chains in Traditional Retailing and E-commerce

Additional packaging

In the e-commerce model, in order to prevent damage to the products during shipment, more packaging materials are used than in the traditional retailing model. To calculate the energy use difference resulting from additional packaging, the usage difference is assumed to be that of extra corrugated boxes, while assuming that the usage difference of other packaging materials is negligible between the two retailing models. As discussed before, the additional packaging of large-sized products in e-commerce model can be neglected.

The type of corrugated box is assumed to be ECT 44 C and the size of the corrugated boxes used for small- and medium-sized products are assumed to be 337 x 322 x 180 mm and 481 x 404 x 389 mm, as suggested by DHL (DHL n.d.). The weight of the corrugated cardboard boxes with a known size can be calculated by using the packaging weight tool (Packaging Calculator 2010): 498 grams (g) for a small box and 1003 g for a medium box. The energy intensity data of the corrugated cardboard is taken from the US Environmental Protection Agency’s (EPA) Waste Action and Reduction Model (WARM) (2012). It is suggested that the 30-year life cycle energy of producing and using one kg of corrugated cardboard material is 28.3 MJ. With these data, the energy for the use of additional material can be calculated with the expression:

\[ E_{add\text{\ packaging}} = e_{corrugated} \times m_{box} \]

Results

Figures 4 and 5 show the results of the previous calculations, and allow for a comparison of the energy consumption differences between the traditional retailing and the e-commerce models.

For traditional retailing, the logistics link with dominant energy usage varies across different types of product: it is the retail store for white goods (61.7%), computer equipment (51.0%), communication devices (68.8%) and health and beauty products (52.0%); the distribution stage for general merchandise and books (69.6% for both); and, customer transport for clothing and shoes (46.1%), and groceries (100%). For almost all types of products, logistics energy usage for traditional retailing is relatively low, ranging from 0.3 MJ (groceries) to 97 MJ (computer equipment), except for white goods, which results in a stunning amount of energy consumption (385.1 MJ) (Fig. 4).
For the e-commerce model, Figure 5 clearly illustrates that the distribution stage is the logistics link with dominant energy usage for almost all types of product, except for grocery products, whose distribution stages are beyond the scope of this comparison. This dominance can be explained by the high-energy intensity of air transportation, and the assumption of the high percentage of air transportation adopted in the distribution stage for the e-commerce model.
A comparison of energy consumption of the logistics chain in the two retailing models is shown in Figure 6. Given the high-energy consumption for heavy product distribution via airfreight, unsurprisingly, the e-commerce model consumes more energy than the traditional retailing model in the logistics chain for all product types, except for white goods. For white goods, the e-commerce model is absolutely in a leading position in terms of energy consumption, as the e-commerce model eliminates the heavy energy-consuming links of the retail store and customer travel in the logistics chain.
Figure 6: Comparing Energy Consumption in the Traditional Retailing and E-commerce Models

Source: Author’s analysis.

Sensitivity Analysis

In this section, two sensitivity analyses are conducted between distribution energy consumption and mixed air and ground transportation, as well as the distribution distance in the e-commerce model.

In the section on Distribution, transportation between the regional logistics centres in the e-commerce model was assumed to be completed via air freight. This results in a mix of 80 per cent air and 20 per cent ground transportation for the distribution link, which can serve as the worst case scenario for the energy consumption calculation. In a more generic case, ground transportation might also be involved in the distribution link between regional logistics centres in the e-commerce model, raising the share of the ground transport, as shown in studies: 12 per cent air and 88 per cent ground (Weber et al. 2008) and 50 per
cent air and 50 per cent ground (Matthews et al. 2001). According to the major Chinese logistics companies, around ten per cent of goods are distributed via air. To reflect a more generic case, a mix of 10 per cent air and 90 per cent ground (with medium-duty truck) transportation is assumed in the e-commerce model for the sensitivity analysis.

Figure 7: Energy Consumption Comparison between Traditional Retailing Model, E-commerce Model with 80% Air and 20% Ground Transportation, and 10% Air and 90% Ground Transportation

Source: Author’s analysis.

As Figure 7 suggests, after decreasing the percentage of air transportation in the e-commerce model, the total e-commerce energy consumption shrinks by about 30 per cent to 60 per cent for most products. Although previously the e-commerce model for clothing and shoes, computer equipment, audio-visual products and health and beauty products consumed significantly more energy than the traditional retailing model, it has now become competitive with the traditional retailing model in terms of energy consumption. For communication devices, the e-commerce model consumes less energy than the traditional retailing model.
Another special feature for retailing in China is that many manufacturers and wholesalers, especially in south-east China, are located nearby to the consumers’ market. So the distribution distance between the wholesaler and the consumer may be smaller than assumed. For the second sensitivity study, a distribution distance of 200 km was assumed for the two retailing models. For this distance, air transportation would be unrealistic and the transportation methods for distribution in the two retailing models were assumed to be the same and were therefore excluded. The results are shown in Figure 8.

Figure 8: Energy Use in Logistics Chains of Traditional Retailing, E-commerce with Distribution Distance of 200 km

Source: Author’s analysis.

As the distribution distance is shortened to 200 km, the advantage of e-commerce is further emphasised for products such as clothing and shoes, computer equipment, communication devices, audio-visual products and health and beauty products. For general merchandise and grocery products, the high-energy consumption for last-mile delivery and additional packaging has prevented the e-commerce model from performing better than the traditional retailing model in terms of energy use. For white goods, the lower energy consumption of the e-commerce model remains.
From these two sensitivity analyses, we see that the energy consumption of the logistics chain in the e-commerce model is dominated by energy consumption in the distribution link, as the percentage of airfreight and the distribution distance can both greatly influence the total energy consumed in the distribution link. The high percentage of air transportation impairs the competiveness of e-commerce against traditional retailing in terms of energy consumption. For a short logistics chain that can be locally completed, however, e-commerce could be more energy-efficient for certain products, as the energy-consuming stages, including the retail store and customer transport, in traditional retailing are replaced with last-mile delivery with less energy consumption. The good performance of e-commerce’s energy-saving features in the short distribution distance scenario suggests the potential of e-commerce to be an effective tool to promote the energy efficiency for local economy activities. By incorporating e-commerce into the local economy structure, further energy efficiency and environmental benefits can be expected.

Conclusion and Further Work

Three logistics chain models were generalised for traditional and e-commerce retailing of products that are often purchased online and the energy consumption at each stage of the logistics chain was assessed against one of the product features in both retailing models. The energy consumptions in both retailing models were compared for each type of product. We found that the distribution stage in the e-commerce model is of the most importance, as its energy consumption is sensitive to a mix of air-ground transportation and distribution distance. Thus the competiveness of e-commerce against traditional retailing in terms of energy consumption is strongly related to conditions at the distribution stage. Varying results for the energy consumption of the two retailing models could be obtained for different types of products.

Further work can be done to improve the descriptive correlation between the product features and the energy consumed throughout the logistics chain, for example by incorporating more product features or specifying more levels to one product feature when calculating the energy consumptions throughout the logistics chain. Also, the methodology proposed in this project can be used to analyse the differences in energy efficiency between the traditional retailing and e-commerce for other countries or regions. As e-commerce increases during the ongoing ICT revolution, understanding these differences will assume even greater importance in the future.
Acknowledgements

I would like to thank Dr Matthew Doolan for his advice, encouragement and help in finishing this paper. I would also like to thank Dr Jane Golley and Ms Yimeng Jiang for reviewing my paper.

Appendix 1

Table 9: Unit Abbreviations

<table>
<thead>
<tr>
<th>Full name</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>gigabyte</td>
<td>GB</td>
</tr>
<tr>
<td>kilogram</td>
<td>kg</td>
</tr>
<tr>
<td>kilometre</td>
<td>km</td>
</tr>
<tr>
<td>kilowatt hour</td>
<td>kWh</td>
</tr>
<tr>
<td>litre</td>
<td>L</td>
</tr>
<tr>
<td>megabyte</td>
<td>MB</td>
</tr>
<tr>
<td>megajoule</td>
<td>MJ</td>
</tr>
<tr>
<td>megawatt hour</td>
<td>MWh</td>
</tr>
<tr>
<td>millimetre</td>
<td>mm</td>
</tr>
<tr>
<td>tonne</td>
<td>t</td>
</tr>
<tr>
<td>watts</td>
<td>W</td>
</tr>
<tr>
<td>yuan (Chinese currency)</td>
<td>¥</td>
</tr>
</tbody>
</table>

Appendix 2

Conversion Between Energy Units

1 kWh = 3.6 MJ

1 MWh = 3600 MJ
Bibliography


Huan Zheng, Yale Wong

Abstract

The need for humanity to move towards a more sustainable future is widely recognised. This transition will require a reduction in the consumption of everyday products, such as sanitary paper. This material stocks and flows analysis explores the use of toilet paper and paper towels at Fenner Hall, a residential community of 507 students at The Australian National University, Canberra. A combination of quantitative and qualitative research methods are used to develop an inputs–processes–outputs model of sanitary paper, consistent with the systems thinking ethos of human ecology. This essay discusses the input flows of the material, before presenting the results of the seven-day data collection process. Daily sanitary paper consumption is estimated to be 70 grams on average per resident, and annual consumption 12.1 tonnes for the entire Fenner Hall community. The study then examines the output flows of toilet paper through the sewerage system, and paper towels through general waste collection. The results of the empirical evidence suggest some gender and temporal variation in the use of sanitary paper, which together with the free nature of this resource, offer potential directions for future research into human behaviour.

Introduction

Given that 70 per cent of the world’s population will live in urban agglomerations by 2050 (United Nations Children’s Fund (UNICEF), 2012), the notion of sustainability being won and lost in cities holds true (Bai, 2010). As such, it is imperative that human activities in urban systems are revisited, through reference to the urban metabolism metaphor, which consists of upstream inputs, processes and downstream output. Material stocks and flows analyses have proven to be an effective outputs of studying the ecosystem service impacts of goods and services, which is important theoretically and practically for the
achievement of broader sustainability objectives. A sector of such importance is the pulp and paper industry, recognised by Sundin et al. (2001) as having far-reaching significance for the triple bottom line objectives of social, economic and environmental sustainability. Not only is paper an indispensable consumable of modern human life, but the global annual consumption of 300 million tonnes has fostered the emergence of global supply chains, according economic benefits to developed and developing countries alike (Verberis et al., 2004: 245). Further, the high system energy requirement in the life cycle of paper contributes to its adverse environmental impacts, resulting in deforestation and greenhouse gas emissions, which have implications for climate change and biodiversity (Martin, 2007, Sundin et al., 2001).

Sundin et al. (2001) categorise paper products into six main grades, according to their end-use purposes: newsprint, graphic papers, corrugating materials, packaging and board, sanitary and household paper, and other paper and board. The focus of this study’s material stocks and flows analysis is on sanitary and household paper products, namely toilet paper and paper towels. This is examined through a case study of sanitary paper consumption at Fenner Hall, a residential community at The Australian National University (ANU), Canberra. The sampled population differs from the wider Australian public, being a community of students, whose behaviour may also be affected by the sanitary paper being provided essentially for free. Firstly, this article explores the background and context of sanitary paper production, use and disposal in the Australian Capital Territory (ACT) region, with reference to its significance in everyday life. Through the presentation of an influence diagram illustrating the variables affecting sanitary paper consumption, this study then overviews the system boundary and methods used in the data collection. Next, the essay examines the inputs–processes–outputs model of sanitary paper with reference to the interdependency of each factor in a systemic approach. The essay examines the results of the analysis, and offers a discussion of the potential weaknesses and limitations of the project, as well as the implications of the research findings on urban sustainability. The systems thinking and holistic approach undertaken in this study pay homage to the complexity of interactions between humans and the urban environment.
Background and Context

From production to use to waste management, the full life cycle of paper products features eight main stages of transformation, as illustrated in Figure 1.

Figure 1: The typical life cycle of paper
Source: Adapted from Sundin et al. (2001).

The sanitary and household paper sector has experienced the highest rate of growth amongst the six categories identified in Sundin et al. (2001) — a global increase of 57 per cent between 1987 and 1996, commanding a market share of 7 per cent of all paper products in 1995 (Sundin et al., 2001: 93). Sanitary paper is unique in that it does not require a separate converting facility, being able to be converted into recyclable material in a paper mill, with strength and softness enhanced through the addition of virgin pulp (Sundin et al., 2001: 93). The toilet paper and paper towel consumption of the 19.9 million urban Australians (Trading Economics, 2012) may be extrapolated from a case study of Fenner Hall. As confirmed by Fenner administration staff (2012), the Fenner Hall community houses 507 students on 20 floors in two towers, served by 61 toilets (an assortment of male, female and unisex bathrooms), 18 kitchenettes and a large shared kitchen on the ground floor. The sanitary paper products examined were produced by Kimberly-Clark and branded Scott for the toilet paper and Kleenex (Optimum Towel) for the paper towels. According to its website, Kimberly-Clark products are certified by the Forest Stewardship...
Council (FSC) as ‘mixed sources’, implying that 50 per cent of the pulp is obtained from FSC-accredited forests, with stringent sustainable management standards and quality control ensured through independent, third-party audits (Kimberly-Clark, 2012a, Kimberly-Clark, 2012b). The toilet paper and paper towel dispensers are refilled every morning by ISS Washroom Services cleaners contracted by Fenner Hall. The journey of the paper products, from their origin as timber in forests and plantations, to their use and disposal as either general waste or sewage is outlined in Figure 2.

**Figure 2: Sanitary paper flows from production to waste management**


The timber for Kimberly-Clark Australia and New Zealand paper products is sourced from forests and plantations in Canada, New Zealand and Brazil (Kimberly-Clark, 2012b). The fibres are transported to the firm’s paper mill in Millicent, South Australia, where toilet paper and paper towels are produced from wood pulp (Kimberly-Clark, 2012b, Sundin et al., 2001). After packaging, they are transported to the Kimberly-Clark warehouse in Fyshwick, Canberra, and then distributed to homes and businesses — in Fenner Hall’s case, to ISS Washroom Services, also located in Fyshwick (ISS Australia, 2012, Kimberly-Clark, 2012b). Staff of the company refill the toilet paper and paper towel dispensers at Fenner Hall on a daily basis. It is expected that toilet paper will be used for personal hygiene following defecation or urination and the paper towels will be used in the kitchens and for drying hands in the bathrooms. It
is at this point that the journeys of the two sanitary paper products diverge. Toilet paper is flushed into the sewerage system, ending up at Canberra’s Lower Molonglo Water Quality Control Centre in West Belconnen, where the solid ‘sewage sludge’ is removed and incinerated in a high temperature furnace. The resulting Agri-Ash is sold to farmers in the region as a soil conditioner (ACTEW Water, 2011). Paper towels are primarily disposed of as general waste, entering the Fenner Hall dumpster and then sent by garbage truck to the Mugga Lane Resource Management Centre in Symonston. In 2010–11, approximately 268,000 tonnes of waste were sent to landfills from the ACT and surrounding towns, such as Queanbeyan (ACT Territory and Municipal Services (TAMS), 2006).

System Boundary and Methods

A system boundary separates the internal components of a system from external entities and, in this study, is defined as Fenner Hall’s physical boundary. The actors within this system comprise of Fenner Hall’s 507 residents, each of whom consumes toilet paper and paper towels based on a combination of factors as depicted by the influence diagram (Fig. 3). The variables are multifarious, and may be categorised as being biological, seasonal, socio-cultural or economic. For instance, the digestive health of residents may influence the frequency of toilet visits, thereby impacting on the use of toilet paper. Similarly, climate and weather affects people’s bodily fluid evaporation, thereby determining their propensity to visit the bathroom. Such seasonal influences also shape food availability, which together with residents’ economic circumstance and cooking styles, dictate the need to use paper towels. Perhaps the most crucial determinant of sanitary paper consumption, however, is the environmental awareness of the residents. This is the only factor that leads residents to make a deliberate and conscious decision regarding their use of toilet paper and paper towels.
Figure 3: Influence diagram depicting variables shaping sanitary paper consumption (text represents variables whilst arrows represent processes)

Source: Author’s own work.

It is beyond the scope of this essay to examine the full spectrum of variables governing the dynamics of the Fenner Hall system. Rather, this study focuses on the key variable of toilet paper and paper towel consumption, using a combination of quantitative and qualitative research methods. Quality control was maintained in the data collection process by following a mixed-method approach to ensure that the research findings are credible. The quantitative section of the study involved data collection over a period of seven days (Monday to Sunday) in the form of measuring the height of the sanitary paper remaining in centimetres (cm) in relation to the vertical capacity of the dispenser, thereby informing the height of paper used (Fig. 4). By comparing this with the paper products original packaging, the number of toilet paper and paper towel sheets could then be determined, and the mass of paper used obtained to provide information on the sanitary paper consumption of Fenner Hall residents and the Australian community more broadly.
Measurements were conducted at midnight so as to account for full sanitary paper consumption during the day before supplies were refilled the next morning. Due to limited human resources, three floors out of ten were randomly chosen from each of the two towers for sampling each night. Every floor consists of three bathrooms — male, female and unisex, equipped with a paper towel dispenser and a toilet paper dispenser in each of two cubicles. A kitchenette with a paper towel dispenser is also present on every floor, primarily used by residents making simple meals with the microwave or toaster. As such, ten measurements were made on each floor, except for the two ground floors, which feature a different design. North Tower ground consists of a single unisex bathroom, while South Tower ground incorporates two male and two female toilets, with neither featuring kitchenettes. Ground floor measurements also included the two male and one female toilet in the foyer, though their use of toilet rolls as opposed to interleaved tissues prevented the sampling of their toilet paper usage. The ground floor kitchens were also surveyed each night, with four out of 13 paper towel dispensers randomly chosen for measurement. The data collection method featured in this study systematically surveyed a sample of 30 per cent for a seven-day period, ensuring an adequate sample size and allowing comparisons to be made between genders, and between weekdays and weekends.
The study conducted also involved a modest qualitative element, involving informal unstructured interviews with the ISS Washroom Services cleaners (2012) and the Fenner Hall administration staff (2012). Comments were sought from the ISS cleaners on the dynamics of toilet paper and paper towel refill processes for each floor every day, as well as for general information about their organisation and their sanitary paper supplier. Interviews with Fenner Hall administration staff involved confirming the number of residents housed, and nuances regarding the contract with ISS Washroom Services. Naturally, the data collection process also encompassed observations, particularly seeking patterns of consumption, users’ attitudes, as well as any material espousing environmental awareness, such as from ANUgreen, the university division working to reduce the collective environmental impacts of students and staff (ANUgreen, 2012).

Data Analysis and Discussion

This section conducts a close study of stocks and flows of toilet paper and paper towels. It first discusses the input flows, namely wood harvesting, raw material processing, pulpwod shipping and paper production, and then examines the stocks and processes which take place within Fenner Hall, with an analysis of the data collection. Finally, measurement of the output flows concerning waste management at the Mugga Lane Resource Management Centre (Symonston) and the Lower Molonglo Water Quality Control Centre (West Belconnen) are presented.

Input Flows

The transportation distance in the life cycle of the Scott toilet paper and Kleenex (Optimal Towel) paper towels found at Fenner Hall were estimated using the Geographic Information System DistanceFromTo (2012). Kimberly-Clark products in Australia and New Zealand began their journey as wood in the forests and plantations of Canada, New Zealand or Brazil (Kimberly-Clark, 2012b). Assuming direct transport to the firm’s paper mill in Millicent, the pulpwood would have travelled an estimated distance of 14,800 kilometres (km), 3000 km and 14,100 km respectively (Kimberly-Clark, 2012a, Kimberly-Clark, 2012b). It is most likely that the subsequent journey of the resulting toilet paper and paper towels took place by road and, given the indirect nature of road networks, 130 per cent of the geodesic distance has been used in this analysis to account for this. As such, the sanitary paper travelled 1137 km from Millicent to the Kimberly-Clark warehouse in Fyshwick (Kimberly-Clark, 2012a, Kimberly-Clark, 2012b). Following this, the paper products are likely to have travelled another kilometre to the ISS Washroom Services facility, also in Fyshwick, before journeying the final 12 km to Fenner Hall in Braddon (ISS

Australia, 2012). In total, the sanitary paper was transported an average distance of 11,780 km from its origin to the point of final use, demonstrating the global scale of supply chain networks.

It is likely that the pulpwood was transported to the paper mill in Millicent by cargo ship, before being sent to Canberra on a container truck, and then moved to Fenner Hall by van, all of which are modes of transport requiring the use of fossil fuels. Fossil fuel use induces social, economic and environmental consequences deemed to be strongly negative, with severe impacts for health, energy security and climate change. Indeed, the transportation sector accounts for 25 per cent of global energy demand and 55 per cent of yearly oil consumption, resulting in direct, indirect and cumulative externalities estimated to cost the world economy more than US$1 trillion per year (Rodrigue et al., 2009). Evidently, the input process for the toilet paper and paper towels at Fenner Hall involves significant transportation and fossil fuel use, with ramifications for the triple bottom line objectives of social, economic and environmental sustainability.

**Stocks and Flows**

The stocks and flows of sanitary paper use at Fenner Hall in terms of the consumption of toilet paper and paper towels constituted the fieldwork component of this study. By extrapolating from the data collected, the total consumption of sanitary paper at Fenner Hall was found to be 12.1 tonnes annually, with an average daily consumption of 70 grams per resident (Table 1). To sustain the flows within the system, Fenner Hall must maintain stocks of approximately 36 kilograms (kg) of sanitary paper daily — 11 kg of toilet paper and 25 kg of paper towels. These significant figures have adverse economic and environmental impacts (see Sundin et al., 2001) and may, in part, be attributable to toilet paper and paper towels being free resources for all 507 residents.

<table>
<thead>
<tr>
<th>Items</th>
<th>Personal consumption</th>
<th>Total consumption at Fenner Hall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily (g)</td>
<td>Annual (kg)</td>
</tr>
<tr>
<td>Toilet paper</td>
<td>21</td>
<td>7.4</td>
</tr>
<tr>
<td>Paper towel</td>
<td>49</td>
<td>16.5</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>24</td>
</tr>
</tbody>
</table>

Source: Author's own work.

The above estimations are based on information provided on the original packaging of the Scott toilet paper and Kleenex (Optimal Towel) paper towels.
studied. Figures on the personal and total consumption of sanitary paper daily and annually were calculated from the height and weight of each paper product, measured in centimetres and grams respectively (Table 2).

**Table 2: Sanitary paper package details**

<table>
<thead>
<tr>
<th>Items</th>
<th>Original packages</th>
<th>Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toilet paper</td>
<td>500 sheets/10 cm height</td>
<td>175 g</td>
</tr>
<tr>
<td>Paper towel</td>
<td>120 sheets/13 cm height</td>
<td>225 g</td>
</tr>
</tbody>
</table>

Source: Author’s own work.

It is worth noting that gender acts as an influential factor determining the daily consumption of sanitary paper. As illustrated in Figure 5, female sanitary paper consumption is on average some 25 to 50 grams higher per day than male.

![Graph showing comparison of daily consumption by gender](image)

**Figure 5: Comparison of average daily consumption of sanitary paper in the floor bathrooms by gender**

Source: Author’s own work.

Clearly, the highest proportion of sanitary paper consumption was found in female bathrooms, followed by the unisex and male bathrooms. It appears that biological differences and gender-based needs are a source of influence on sanitary paper consumption, though a more rigorous statistical analysis is
required to verify this. Indeed, this gendered paper consumption pattern may also be the result of variations in personal hygiene habits and environmental consciousness. Evidently, only the latter two factors are manageable by policy makers, though their importance in the mix of variables that account for gender consumption differences remains unclear.

A day-by-day comparison was also drawn to display the variation of toilet paper and paper towel consumption over a period of one week (Figure 6).

![Figure 6: Day-by-day breakdown of sanitary paper usage](image)

Source: Author’s own work.

As depicted by the graph, the highest consumption of sanitary paper was found on Monday, with a usage of nearly 12 kg, while Friday constituted the lowest consumption, it being some 30 per cent lower than that of Monday. Indeed, the total daily sanitary paper consumption remained high from Monday to Wednesday, dropping on Thursday and Friday, before rebounding over the weekend. Based on the experiences and observations made by the researchers, who also reside at Fenner Hall, such variations may be attributable to the majority of residents maintaining a regular study and cooking regime from Monday to Wednesday. Towards the end of the working week, many residents have fewer, if any, contact hours at the university and may dine outside and party from the evenings until the early mornings. Such patterns reduce the frequency of kitchen and bathroom visits, which is particularly pertinent towards the evening. Over the weekend, some residents may stay at Fenner Hall and maintain the usual
frequency of cooking and toilet usage, whilst other residents may take leave from Fenner Hall for one or two days, thereby explaining the gradual rebound of the sanitary paper consumption depicted in Figure 6.

As the influence diagram (Fig. 3) previously demonstrated, a multitude of biological, seasonal, socio-cultural and economic variables interact with one another to determine the flows in the system; i.e. the consumption of toilet paper and paper towels. The stocks of the daily 36 tonnes of sanitary paper consumption, and by extension, the annual 12.1 tonnes, are shaped by both internal and external factors. As such, it is imperative to acknowledge the interconnectedness and complexity of the studied system, recognising the need for systems thinking to change the complex stocks and flows of sanitary paper consumption.

In general, the stocks and the consumption processes of toilet paper and paper towels at Fenner Hall remain significant in spite of the mild variations between gender and the days of the week. It is necessary to note that both the 70 grams per resident and the total 36 kg of daily sanitary paper consumption are considerable rates of flow. As such, there is scope for a more detailed investigation into Fenner Hall residents’ sanitary paper consumption behaviours and patterns. Having a more nuanced understanding of such stocks and processes could inform policies to reduce toilet paper and paper towel use at Fenner Hall, with implications for the Fenner Hall budget and the environment more broadly.

Output Flows

Given that toilet paper is used for personal hygiene following defecation or urination, and paper towels are for kitchen use and for drying hands, their methods of disposal are different, resulting in a divergence in the journeys made by these two sanitary paper products. After being flushed from the toilet bowl, toilet paper then enters the sewerage system, being conveyed by wastewater an estimated distance of at least 19 km to arrive at the Lower Molonglo Water Quality Control Centre in West Belconnen (ACTEW Water, 2011). This sewage is then treated, with the liquid effluent either reused for public garden irrigation, or reentering the Molonglo River, flowing downstream into the Murrumbidgee outside the city. The water then joins the Murray–Darling River, before emptying into the ocean at Lake Alexandrina in South Australia. The majority of toilet paper, however, will remain in solid form and be removed as ‘sewage sludge’, incinerated and then sold to farmers as fertiliser (ACTEW Water, 2011).

Used paper towels are disposed of in the waste bins found in the bathrooms, kitchenettes and the ground floor kitchen at Fenner Hall. ISS Washroom Services cleaners collect these and move them to the Fenner Hall dumpster, which is located behind the main kitchens. This general waste is removed by
waste collectors in garbage trucks, and travels at least 16 km from Braddon to the Mugga Lane Resource Management Centre in Symonston (ACT TAMS, 2006). The waste may undergo some processing, but is ultimately compacted and buried, decomposing slowly due to a lack of oxygen and moisture. As such, the annual consumption of paper towels at Fenner Hall of some 8.4 tonnes has a significant environmental impact, with consequences for surface water, groundwater, flora, fauna, air quality and noise (South Australian Environment Protection Authority, 2007). The environmental ramifications of Fenner Hall’s 3.7 tonnes of annual toilet paper consumption are, however, somewhat mitigated by the reuse of ‘sewage sludge’ in farming (ACTEW Water, 2011), though the incineration process does generate emissions and toxins. Fine particulate matters produced near incinerators have been found to increase morbidity and mortality through cancer, birth defects and heart disease (Thomson and Anthony, 2005).

**Implications for Urban Australia**

Although this material stocks and flows analysis was conducted through a case study of Fenner Hall, its findings regarding toilet paper and paper towel consumption may be used to obtain an order of magnitude estimate encompassing all 19.9 million urban Australians (Trading Economics, 2012). Drawing on the figures in Table 1, it has been found that urban Australians consume a total of 1400 tonnes of sanitary paper daily and 511,000 tonnes annually. This number was devised through simple numeric extrapolation, on the assumption that Fenner Hall residents are roughly typical in their consumption patterns. Such large figures embody the realities of the modern growth economy, based on perpetual increases in the production and consumption of goods and services, bringing to light the social, economic and environmental consequences of the market system. To alter the material flows of products in the Australian economy, governments may implement market-based policy instruments or legislate regulatory measures to internalise negative externalities. In recent times, however, there have been renewed calls to foster cultural change, managing demand by raising environmental awareness, aided by actors like governments, non-governmental organisations and community groups. In the case of Fenner Hall, ANUgreen and the Fenner Hall Administration Office have attempted to build residents’ environmental consciousness through heightened awareness of their electricity and water usage.

There is scope to expand this program to encompass the nuances of toilet paper and paper towel consumption by targeting the potential gender-based consumption patterns and residents’ personal habits. For instance, an inter-floor competition could be held, with a prize awarded at the end of the semester to the floor using the least amount of sanitary paper. Naturally, the feasibility of
such exercise is dependent on human resources and the propensity for residents to sabotage the competition. This final notion regarding the importance of collectivism in solving complex environmental problems also grounds the study conducted with theoretical significance. As toilet paper and paper towels are provided free of charge at Fenner Hall, there is reason to believe that their consumption may suffer from the ‘tragedy of the commons’ (Hardin et al., 1973). As non-excludable, but rival products, each resident has an incentive to consume the sanitary paper as rapidly as possible before it is used up. One way to counter this tendency would be to offer sanitary paper at a fee. This is particularly pertinent for paper towels in the ground floor kitchen, which was found to have been frequently left with little or no stocks at the end of each day. Paper towel is, however, a discretionary commodity, being easily replaceable with the alternative of cloth towels and is hence an easier target for instituting social change. Nevertheless, potential future projects may involve interdisciplinary research with economists and psychologists to further understand the dynamics of human behaviour. As such, this material stocks and flows analysis has served not only a practical, but also theoretical significance, offering itself as a springboard for further research projects.

Limitations

The study faced potential methodological weaknesses during fieldwork and data analysis, including both random and systematic errors. Firstly, human error may have resulted in inaccurate measurements during the data collection process — a particular concern being parallax error when reading off the height of sanitary paper used. Fatigue may have contributed to transcription errors, which would have impacted on the reliability of the results, with ramifications for data analysis and the implications drawn.

Such challenges are coupled with issues in statistical inference, namely in making generalisations about the Fenner Hall community and urban Australians more broadly. In each of the seven days, six of the 20 floors and four of the 13 dispensers on the ground floor kitchen were surveyed, raising the question of whether our sample size is sufficient to be representative. In the same way, it is unknown whether a consecutive seven-day study period in early November 2012 can illustrate the average sanitary paper consumption of Fenner Hall residents, particularly given that the study was conducted during students’ preparation period for examinations. They may have had a lower propensity to cook, resulting in skewed paper towel consumption figures. As noted in the influence diagram (Fig. 3), there is a seasonal influence upon sanitary paper consumption, as average daily temperatures can impact upon residents’ bodily fluid evaporation, thereby altering their toilet visitation patterns. It must also be
noted that as Fenner Hall residents are from interstate or overseas, the majority are in residence for a period of nine months during the academic year only, and so the aforementioned yearly sanitary paper consumption figures may not have represented the reality consumed. Finally, the study did not account for residents who chose to purchase their own sanitary paper or use cloth towels as opposed to paper towels for cooking purposes, such as drying dishes or mopping up spills.

These random errors are also coupled with potential systematic errors. Indeed, the measured height of sanitary paper, conducted at midnight, may not have corresponded to the actual use of toilet paper and paper towels. According to the interview with ISS Washroom Services cleaners (2012), they refilled the dispensers only in the mid-morning, offering a window for residents to use sanitary paper in the early morning, which would not have been captured in this study. Although this limitation was made note of during the study’s design phase, it appeared inevitable due to practicality and the time constraints of the researchers in the data collection process.

Another point of contention arises from the use of the toilet paper and paper towel consumption figures of Fenner Hall residents to make generalisations about the broader urban Australian community. Much like the notion of ‘aggregation bias’ as described in Ewing and Cervero (2010), making inferences about the aggregate from the individual may be problematic for a range of reasons. One is that the lifestyle choice of university students may differ from that of typical Australians, with people from the targeted age group perhaps being more prone to wasteful consumption. It has been argued, however, that the younger generation is also more environmentally conscious, and hence the impact of the studied demographic on the aggregated results remains unclear. Moreover, some 50 per cent of Fenner Hall residents are international students, with cultural differences having the potential to shape their sanitary paper consumption, thereby contributing to the complexity of this exercise.

**Conclusion**

This material stocks and flows analysis examines the consumption of toilet paper and paper towels in the urban context of the Fenner Hall residential college. This essay presents the input flows of the material, namely wood harvesting, raw material processing, pulpwod shipping and paper production, before conducting a seven-day data collection process to estimate daily sanitary paper consumption to be 70 grams per resident and annual consumption to be 12.1 tonnes for the entire residence. The study then examines the output flows of toilet paper through the sewerage system, and paper towels through general waste collection.
This essay offers a glimpse of the interdependence of the biological, seasonal, socio-cultural and economic variables, thereby supporting the nuances of the influence diagram presented (Fig. 3), emphasising the importance of systems thinking. The essay acknowledges humanity’s need to alter flows, and proposes economic, political and cultural instruments to bring about this change. It then presents the potential weaknesses and limitations of the research conducted, giving grounds for review and improvement. In conclusion, the urban system is complex and interconnected, due to it being of a heavily human-modified nature. Australians’ high use of sanitary paper stems from it being seen as an indispensable part of modern daily life, which together with the consumption of other such goods and services, results in the intensive utilisation of ecosystem services. Only a transformational change which closes the loop in the life cycle of material stocks and flows by reducing materials can enrich our cities, moving them towards the triple bottom line objectives of social, economic and environmental sustainability.

Acknowledgements

We would like to thank Dr Rob Dyball for conceptual assistance during the course of this project and Dr Barry Newell for his constructive advice on the crucial revisions made to this essay. The useful suggestions from Professor Steve Dovers and Brendan Moloney are also very much appreciated. Finally, we thank all 507 residents of Fenner Hall for putting up with our presence on many late nights and for intruding on their bathroom activities.
Appendix: Summary of Raw Data from Day 1 to Day 7

<table>
<thead>
<tr>
<th>Day</th>
<th>Gender</th>
<th>Toilet Paper</th>
<th>Kitchen Paper</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Monday</strong></td>
<td>Female</td>
<td>85</td>
<td>104.5</td>
<td>Sum</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>86</td>
<td>96</td>
<td>Toilet Paper 252.5</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>135</td>
<td></td>
<td>Kitchen Paper 434</td>
</tr>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tuesday</strong></td>
<td>Female</td>
<td>75 (-1)</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>72</td>
<td>110.5</td>
<td>Toilet Paper 163</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>136</td>
<td></td>
<td>Kitchen Paper 497.5</td>
</tr>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Wednesday</strong></td>
<td>Female</td>
<td>70 (+1)</td>
<td>122.5</td>
<td>Sum</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>37 (+2)</td>
<td>142</td>
<td>Toilet Paper 183.5</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>136</td>
<td></td>
<td>Kitchen Paper 494.5</td>
</tr>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Thursday</strong></td>
<td>Female</td>
<td>59 (-1)</td>
<td>80.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>79.5</td>
<td>75.5</td>
<td>Toilet Paper 168.5</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>126</td>
<td></td>
<td>Kitchen Paper 385.5</td>
</tr>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Friday</strong></td>
<td>Female</td>
<td>77 (-1)</td>
<td>103</td>
<td>Sum</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>56</td>
<td>71</td>
<td>Toilet Paper 123</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>134</td>
<td></td>
<td>Kitchen Paper 400.5</td>
</tr>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Saturday</strong></td>
<td>Female</td>
<td>88</td>
<td>91.5</td>
<td>Sum</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>63.5</td>
<td>64.5</td>
<td>Toilet Paper 222.5</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>132</td>
<td></td>
<td>Kitchen Paper 382.5</td>
</tr>
<tr>
<td><strong>Unisex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sunday</strong></td>
<td>Female</td>
<td>94</td>
<td>97.5</td>
<td>Sum</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>61.5 (-1)</td>
<td>94</td>
<td>Toilet Paper 199.5</td>
</tr>
<tr>
<td></td>
<td>Kitchen</td>
<td>136</td>
<td></td>
<td>Kitchen Paper 421</td>
</tr>
</tbody>
</table>
Bibliography


ISS Washroom Services Cleaners (2012), *pers. comm.*, Relevant information about sanitary paper consumption in Fenner Hall. Canberra.


