The ANU Undergraduate Research Journal

Julia Brown

Volume Nine, 2018

Contents

Foreword

It gives me great pleasure to introduce the ninth volume of the *ANU Undergraduate Research Journal*.

The University takes huge pride in the calibre of the students who choose to study here, and this journal gives us the opportunity to showcase some of the best essays submitted in recent years.

ANU was established by the Australian Government more than 70 years ago with a special charter – to give the nation a world-class university.

To that end we offer the best possible research and education opportunities to our staff and students, who come to ANU from many countries around the world.

Our world-leading researchers undertake research addressing significant national and global challenges, and they are frequently recognised for this work with awards and fellowships. Our expertise informs policymakers and influences the shape of the nation. Many of our students go on to have highly influential and successful careers.

This is the environment that stimulates our students and it is wonderful to see that reflected in these pages.

The range and depth of these essays is impressive and gives us great confidence in the people who will graduate with ANU degrees in the future.

I congratulate the authors and editor of the 2019 journal on their outstanding work and contribution to advancing Australia's research cause.

Professor Brian P. Schmidt AC

Vice-Chancellor and President

The Australian National University

Editorial note

It is 2019. Research communication efforts are perhaps more critical than ever. The undergraduate students at ANU continue to have meaningful things to say. Some wish for *you,* too, to engage with what they learn and reason about as their critical thinking skills begin to blossom. The written contributions in the *ANU Undergraduate Research Journal* (AURJ) are an offering to improve preliminary debate and reflection on things that make up our world.

From global affairs to Australian politics and law, to scientific and humanities explorations, to analysing artistic expression, there is something to be learned from each article in Volume 9 of the AURJ. This cross-disciplinary collection speaks to the strength of the academic community at ANU to include different priorities and fields of enquiry. These articles were selected from a competitive pool of High Distinction essays, and then developed even further into pieces to be appreciated by a wider audience.

It is not always easy to go back and alter a piece that has already consumed so many hours of work. Each student rose to that challenge and reminded me how research and writing can be rigorous at all levels of study. Indeed, these articles speak to the great potential to be found in our undergraduate students as our future thought leaders and communicators.

I edited this volume with the benefit of my work as a Learning Adviser at the ANU Academic Skills centre. Putting it together gave me the opportunity to see the instruction of these academic skills right through to production point. For instance, I was delighted when Nicholas Blood, author of 'Recent Progress Towards Achieving an International Plastics Convention', humbly wrote to me in the margins of his draft to confirm the structure we had discussed: 'Background: The plastic situation is bad. Debate: Even an undergraduate like me can point to ways in which it's likely even worse than the latest research suggests. Argument: To solve this we need to prioritise it globally, but we also need to encourage other approaches at the same time'. Beautiful!

Needless to say, Blood and all authors in this volume have worked hard to assert their reasoned ideas. Some articles evolved significantly from the already top-quality standard that earned their selection, and this is a reality of academic practice and integrity – to keep refining while incorporating other literature until readers can simply focus on the topic, rather than becoming distracted by writing missteps. In this way, the process of having an article published in the AURJ really does help to prepare our undergraduate students for future academic or writing pursuits.

I would like to give enormous thanks to all authors for their willingness and commitment to participating in the hard work behind this present volume. Much gratitude also goes to the copy-editor Beth Battrick, and to my colleagues at both Academic Skills and the Student Experience and Career Development Office at ANU for their support throughout the editing process.

May this volume inspire many others!

Dr. Julia Brown

AURJ Volume 9 Editor

About the authors

James Atkinson

James is a final-year student of a Bachelor of Development Studies and a Bachelor of Arts (political science major). He is interested in the relationship between media and politics, and how it shapes policy and development. He recently returned from a short course in Myanmar with the College of Asia and the Pacific.

Sarah Barrie

Sarah has recently graduated with a Bachelor of Laws (First Class Honours) and a Bachelor of Central Asian and Middle Eastern Studies from ANU. She has a particular research interest in the Middle East, international humanitarian law and interdisciplinary legal studies.

Nicholas Blood

Nick is in his third year of a Bachelor of Interdisciplinary Studies (Sustainability). Much of his learning so far has been relayed into articles on sustainability written for the ANU student newspaper *Woroni*, and the national student newspaper *Et Cetera.* Nick is the 2019 ANU Students' Association Environment Officer and hopes to further apply his learning towards a more sustainable campus and community. Nick's article is a response to a 'Grand Challenge' video on the problems of plastic use, including product packaging. The video was presented to students are part of the course ENVS 2013 – Society and Environmental Change.

Jessica Elliott

Jessica graduated with a Bachelor of Arts and a Bachelor of Laws, Honours (First Class) in 2018. In 2019 she is the Tipstaff to the Chief Justice of New South Wales. She was an Academic Observer to the Australian Mission at the United Nations, a paralegal at Aboriginal Legal Service and Summer Clerk at King & Wood Mallesons. She was a Senior Student Editor of the *Federal Law Review* and *Australian Year Book of International Law*.

Tom Goodwin

Born and raised in Melbourne, I was fortunate enough to receive a Tuckwell Scholarship to study a Bachelor of Laws (Honours) and a Bachelor of Politics, Philosophy and Economics at ANU. I am and will always remain thankful for the opportunities that have been given to me. Inherent in my gratitude

for the start to life Melbourne has given me is an appreciation for the immense formative influence that gold has had on the cultural, economic and political prosperity of my hometown; an influence that I am proud to have investigated in my article.

Georgie Juszcyk

Georgie is a third-year Law/International Security Studies student. Georgie currently works as a Research Assistant for Associate Professors Sarah Heathcote and Asmi Woods, and as an Administrative Assistant at Lexbridge Lawyers. Georgie's primary research interests include national security law, international law, constitutional law and administrative law.

Rebecca Kriesler

Rebecca is currently completing her Bachelor of Laws (Honours) and Bachelor of Arts degrees, majoring in history and international relations. She has focused her studies around researching social justice issues and the persecution of minority groups in the law. She is doing ongoing research with the CEO of Legal Aid ACT to publish a legal guide assisting self-represented litigants to navigate the criminal justice system. She is passionate about ensuring the legal system is accessible for all and aims to pursue this into the future.

Clare Langley

Clare is currently completing a Bachelor of Medical Science at ANU. Clare's major focus has been on biology; however this paper was prepared as a part of a Population Health course.

Kida Lin

I am currently an Honours student in philosophy at ANU. My research interests are in normative ethics. I have recently completed a Bachelor of Arts in philosophy and government/international relations at the University of Sydney, where I received the Lithgow Scholarship No. IV for best performance in senior year of philosophy, and the Francis E Snare Memorial Prize for best performance in moral and political philosophy. In my free time, I enjoy debating and writing. I am a reporter for *ANU Observer*, and I have contributed to *Woroni* and *Honi Soit*. I was also a top 10 best speaker at the Australian Intervarsity Debating Championships in 2018.

Joshua Ling

Joshua graduated from ANU in law and international relations in 2018. In his time at ANU, Joshua was an active participant in various law student competitions, was awarded the Ann Downer Memorial Prize and the Sarah Avery Prize, and was a student editor for the *Federal Law Review* and the *Australian*

Year Book of International Law, both published by the ANU College of Law. He looks forward to exploring a career in dispute resolution.

Alex Lombard

Alex is currently studying philosophy, mathematics, linguistics and Ancient Greek. This paper was originally submitted as an essay for the course Philosophy of the Cosmos (PHIL2042/MATH1042), an introductory course on the philosophy of science and cosmology jointly offered by the School of Philosophy and the Mathematical Sciences Institute.

Annika Morling

I am a third-year PhB student majoring in English. I am commencing my Honours project at the beginning of 2019, in which I plan to look at gothic cinema. My main interest is in film and I am also interested in the gothic genre and nineteenth-century British literature and society.

Julian Moss

Julian is a final-year undergraduate law student at ANU. To support himself, he works part-time as a paralegal. He is passionate about public law, in particular the legality of restrictions on voting rights. In his spare time, he enjoys gardening and bushwalking on Canberra's many great trails.

Dana Royle

Dana is a fourth-year student at ANU, studying a double degree in a Bachelor of Criminology and Science (Psychology). Dana has had a lifelong passion for all things creative and loves reading crime and mystery novels, driving her research interests at ANU.

Yee Seng Tay

Tay is currently completing a Bachelor of Arts degree, majoring in international relations and geography, and will be pursuing his Honours coursework at the Fenner School of Environment & Society in July 2019. His primary area of academic interest lies in environmental policy, biodiversity conservation and global climate change cooperation.

Chenghao Yu

Chenghao is a third-year Bachelor of Science (Psychology) student at ANU. He was awarded the Judith A. Slee Prize for Scientific Writing in Psychology in 2017, attended the IARU (International Alliance of Research Universities) Global Summer Program in 2018 as an ANU Grant recipient and is currently

working as vacation scholar at the Monash Accident Research Centre. He is intrigued by new explorations and applications in the field of behavioural science as he believes it ultimately serves the purpose of empowering more people to have a better life.

Jiahuan Zhang

Jiahuan graduated from ANU with a Bachelor of Arts degree. Her major is linguistics with research interest in the interface study of second language acquisition and language testing. She has finished two pilot projects in her final year concerning Chinese-L2 learners' acquisition of classifiers, which is also the thesis topic for her current Honours year. She is also interested in teaching Chinese as a second language. As early in 2015 and 2016, she gained teaching experience from working as a Chinese language assistant as well as a Chinese culture teacher for the worldwide American government–sponsored summer program National Security Language Initiative for Youth (NSLI-Y) in China, and has received awards for her excellent teaching.

About the editor

Julia Brown

Julia is a Visiting Fellow at the School or Archaeology and Anthropology at the ANU, where she recently completed her PhD in social and medical anthropology. Her ethnographic research in Australia and the UK focused on the lived experiences of clozapine-treated schizophrenia patients and their clinical caregivers. Julia hopes to continue working at the intersection between critical anthropology, public policy and the social studies of science, technology and medicine. She has several publications in reputable international journals and has both tertiary teaching and research assistant work experience, including government public health policy experience. She also co-founded the ANU-supported anthropology social engagement project (blog and podcast) *The Familiar Strange*. Julia is passionate about making academia more accessible and she worked as a Learning Adviser at the ANU Academic Skills centre throughout her PhD candidature.

Cover art

Belle Palmer

*Symbiosi*s, 2018

Interactive installation: bicycle forks, steel, bamboo, hair, string

150 x 500 x 50 cm

Photographer: David Lindesay

Belle won the Peter Karmel Award for this piece, which she completed as part of her Bachelor of Visual Arts (Honours).

Recent progress towards achieving an international plastics convention

NICHOLAS BLOOD

Abstract

This paper identifies that plastic is not only a severe environmental problem, but also a global existential risk that needs greater prioritising in global policy. While issues of macro-scale pollution are increasingly well-known, the role of plastic in other areas such as micro-scale contamination, and the strain plastic production places on global resource limits, enjoys less attention. Focusing on recent data and new or emerging developments, I examine the extent to which a binding international convention on plastics can be considered critical in addressing the many problems caused by plastic and plastic packaging. I argue for the need to understand the problems of plastic holistically and solve them in the same way; by recognising that treaties and conventions are just one critical step among others. Business-led innovations towards circular economies and other emergent, self-organising movements are of equal importance. These other developments represent both near-term solutions, achievable before a convention will be realised, and the precursory steps necessary for creating a binding, international plastics convention.

Introduction

The human use of plastic poses an extreme environmental problem. Approximately 70 per cent of Earth's surface environment is aquatic (US Geological Survey, 2016), and in those plentiful oceans plastic and plastic packaging creates the most significant and persistent pollution. Under a business-as-usual scenario, by 2050 the oceans will contain more plastic than fish (Ellen MacArthur Foundation & World Economic Forum, 2017). Recent research suggests that only 13 per cent of our marine environments remain as true wildernesses untouched by human impacts (Jones et al., 2018). Alarmingly, however, this figure entirely fails to account for microplastics (fragments of plastic <5 mm long), the quantities and locations of which are unknown, and which are an emerging field of study (National Oceanic and Atmospheric Administration, 2018).

This paper argues that plastic is not only a severe environmental problem, but also a global existential risk that needs greater prioritising in global policy. Further arguments about the extent to which global policy should be prioritised are explored within the case study of a proposed binding, international convention on plastics, which serves as an example of a global policy response.

Assessing policy first requires a holistic understanding of the issues it will affect, however. This paper begins by examining plastic production and use across a range of contexts to provide a broad assessment of its impacts and problems. I then examine to what extent a binding international convention on plastic use can be considered as the most important next step in combatting these issues, highlighting numerous barriers to a convention's success that suggest a more multifaceted approach may be ideal. I conclude by arguing that although important, a binding convention requires other precursory steps to occur first, and that an ideal approach to this issue is one that likely incorporates multiple solutions simultaneously.

Understanding the plastics problem holistically

Recent research suggests microplastic contamination is more widespread than anticipated, highlighting how deep and pervasive the problem of plastic pollution has become. One study sampled salt brands around the world and found that 92 per cent of them contain microplastics (Kim et al., 2018). This study follows the slightly earlier finding of widespread global contamination in a range of consumer products, including 81 per cent of sampled tap water (Kosuth et al., 2018). Macro-sized plastics meanwhile contribute to large-scale pollution such as the Great Pacific Garbage Patch, one of the ocean's largest accumulations of plastic, estimated at 1.6 million square kilometres in size – twice the area of New South Wales (Lebreton et al., 2018).

Plastic packaging is an issue of growing popularity, and of significance within the broader question of plastic use. Over a quarter of all plastic is used for packaging and a staggering 86 per cent of it will not be recycled, resulting in the loss of USD 80–120 billion annually in value, and creating a conservatively estimated USD 40 billion in environmental damage (Ellen MacArthur Foundation & World Economic Forum, 2017; World Economic Forum, Ellen MacArthur Foundation & McKinsey & Company, 2016). A narrow focus on plastic packaging alone, however, is not enough to solve the problems of plastic use, since it comprises only a quarter of all use. Additionally, we often cannot tell one type of plastic pollution from another. In the case of microplastic contamination, it is 'very hard or even impossible' (Veiga et al., 2016, p. 9) to tell whether the plastic originated as packaging or in some other form.

With such low recycling rates, it remains likely that packaging contributes significantly to global plastic pollution. This likely includes microplastic contamination. The often small and lightweight properties of packaging make it 'particularly prone to escaping collection systems and ending up in the natural environment, especially in emerging economies where most of the leakage occurs' (Ellen MacArthur Foundation & World Economic Forum, 2017, p. 17). While packaging is clearly a significant area of concern, it must be seen as part of a larger issue involving variables which are still not fully understood, such as microplastic contamination.

Beyond the profound and pervasive degradation of our oceans and food sources, the necessity of fossil fuel feedstocks in a range of plastic production, and plastic's contribution to climate change through greenhouse gas (GHG) emissions are equally large concerns. Under a business-as-usual scenario, by

2050 plastic production will require 20 per cent of global oil production and consume 15 per cent of the world's annual carbon budget (Ellen MacArthur Foundation & World Economic Forum, 2017).

Plastic is polluting our food and drinking water, our bodies, our oceans and land. It is straining uniquely important resources like oil and consuming far too much of our carbon budget (Ellen MacArthur Foundation & World Economic Forum, 2017). The threats posed by plastic use are so severe they can be considered cumulatively as an existential risk to both our species and our planet. Existential risks are an overriding sustainability concern, and arguably therefore a global priority (Bostrom, 2013). In recognition of this and other factors, some have proposed the creation of a global convention on plastics use.

Towards a plastics convention

A good example of a global policy response to plastic comes from two academics at the Heinrich Böll Foundation (HBF), a progressive German think tank. They are key among those who have argued the need for an international plastics convention to address the problems of plastic:

> A global convention that tackles plastic pollution where it originates, fosters innovation for more sustainable plastics, and supports countries in enhancing their domestic waste collection and recycling systems. It is the necessary next step and should have priority, rather than focusing on the Sisyphean task of cleaning up entire oceans while millions of tons of plastic waste keep streaming into them. (Simon & Schulte, 2017, p. 7)

Their characterisation of clean-up projects as Sisyphean is perhaps inaccurate, as is the narrow language in their report which gives priority to a convention and treats it as '*the* necessary next step' (emphasis added). The Ocean Cleanup project (2018), for example, is just one engineering solution that offers a potential 50 per cent reduction in the Great Pacific Garbage Patch's size over just five years. Solutions such as this that operate (for now) outside the realm of convention, deserve better consideration. Indeed, the authors' misstep here emphasises the broader need for the kind of multi-solution approach given to 'wicked problems' like plastic; one that does not ultimately prioritise a single solution (Harris et al., 2010). This point is evidenced elsewhere in their focus on the sources of waste and improving recovery rates of plastic. While important, 30 per cent of all plastic still cannot currently be recycled (Ellen MacArthur Foundation & World Economic Forum, 2017).

The difficulty of recycling all plastics represents an immediate and obvious barrier to increasing recycling rates and limiting plastic pollution. As the report admits, innovation will be required. Their examination of innovations focuses on the weaknesses of so-called 'sustainable alternatives' to plastic, such as downgrading (replacing plastic with paper, glass, etc), a suggested 'innovation' that would increase environmental costs from $139 billion to $533 billion annually (Trucost, 2016).

Notably, while the HBF report identifies a range of workable innovations, these include many already being undertaken by industry bodies such the American Chemistry Council (ACC), which represents

and works with a range of important private sector stakeholders such as the primary producers of plastic and their transnational corporate consumers. The ACC-led Materials Recovery for the Future (MRRF) group, for example, involves some of the world's largest plastic-using businesses and plastic producers, including Procter & Gamble, Target, Dow Chemical, PepsiCo, Nestlé USA, The American Plastics Industry Association, and Chevron (Materials Recovery for the Future, 2018). They are *already* working in concert, focusing on recycling and innovations in sustainable plastic alternatives, just as the HBF report desires.

As the above examples indicate, much progress is likely to be realised before a global policy in the shape of a convention begins to form. As the authors again admit:

> a plastics convention is not assumed to replace all other existing efforts, but to complement them: To establish a legally binding roof on top of the many strategies, action plans, and partnerships out there. (Simon & Schulte, 2017, p. 7)

This approach is even more evident in their framework of five essential features or 'pillars' required for it to succeed (Figure 1).

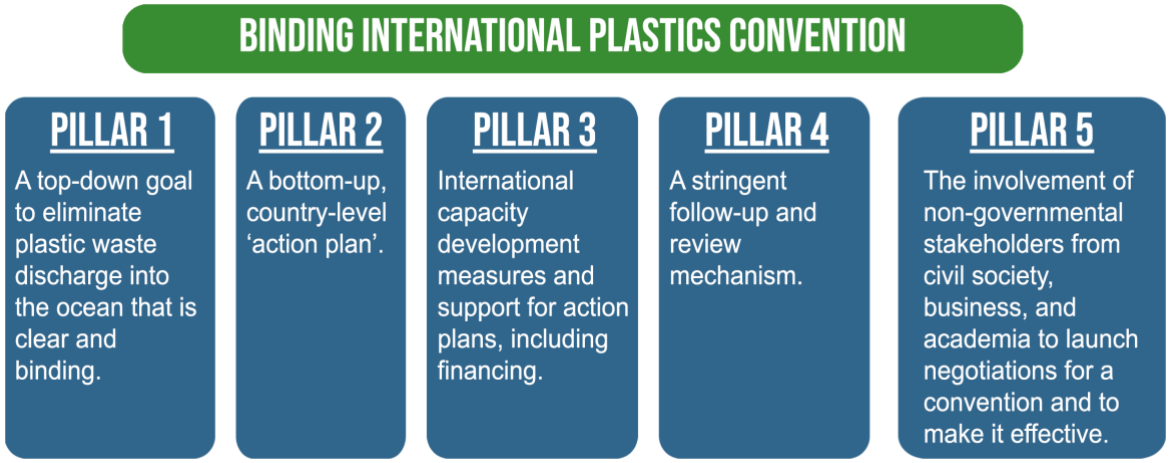

Figure 1: Five pillars of success.

Source: Adapted from Simon & Schulte, 2017, p. 9.

Notably, involvement from other bodies is deemed as critical to the *formation* of the convention as well, and not just its overall effectiveness (Pillar 5). This pillar represents perhaps the first step, and one currently underway. This part of their blueprint has proven prescient only a year later.

The Ellen MacArthur Foundation (EMF) has long studied the issue of plastics and plastic packaging, and in late 2018 began the work of establishing exactly the kind of body described in HBF's approach. This combines academia, civil society, and business, and can help launch negotiations for a convention with the UN. They are building 'a new coalition of businesses and governments united behind a world-leading set of circular economy commitments tackling plastics waste' (Ellen MacArthur Foundation, 2018). Like the ACC-led group mentioned earlier (MRRF) their business partners include primary

producers, as well as large transnational corporate consumers of plastic such as Coke, Walmart, Unilever, Pepsi Co, L'Oréal, Evian, and Nestlé.

Importantly, the EMF is also working with Erik Solheim, UN Environment Executive Director, and the UN Environment body more broadly, hoping to build on recent agreements such as the G7 Ocean Plastics Charter (Group of Seven, 2018), which failed to secure even non-binding commitments from the US and Japan. The involvement of key businesses, governments and the UN gives us perhaps the best chance yet for a global and binding convention. However, there are some barriers to a truly inclusive, cross-institutional and multi-solution approach.

Barriers to a convention

The EMF's coalition is largely guided by their latest research, outlined in the *New Plastics Economy* report (Ellen MacArthur Foundation & World Economic Forum, 2017). This paper has already come under some criticism from other groups. Unfortunately, this includes the ACC-led MRRF mentioned earlier, arguably a key ally going forward. Disagreements between these groups are hardly fundamental, and instead relate to what areas to focus efforts and resources on (Russell, 2017). The early split is nonetheless concerning if a truly inclusive, multi-stakeholder coalition is to be created. The MRRF, with its focus on recovery and recycling, is also arguably close in conceptual alignment with the HBF plan – something that could help in the creation of a convention.

Another deeper issue in regard to the regulation of plastics is that of economic growth. As Alier (2009) states bluntly: 'economic growth is not compatible with environmental sustainability' (p. 1099). It can be argued that a plastic convention alone still cannot solve the underlying issues of plastic production, because they are inherently tied to global consumption trends that are, in turn, a product of the capitalist, neoliberal model of endless economic growth. Only once in the past 50 years has plastic production abated, and that was following the 2008 Global Financial Crisis (Figure 2).

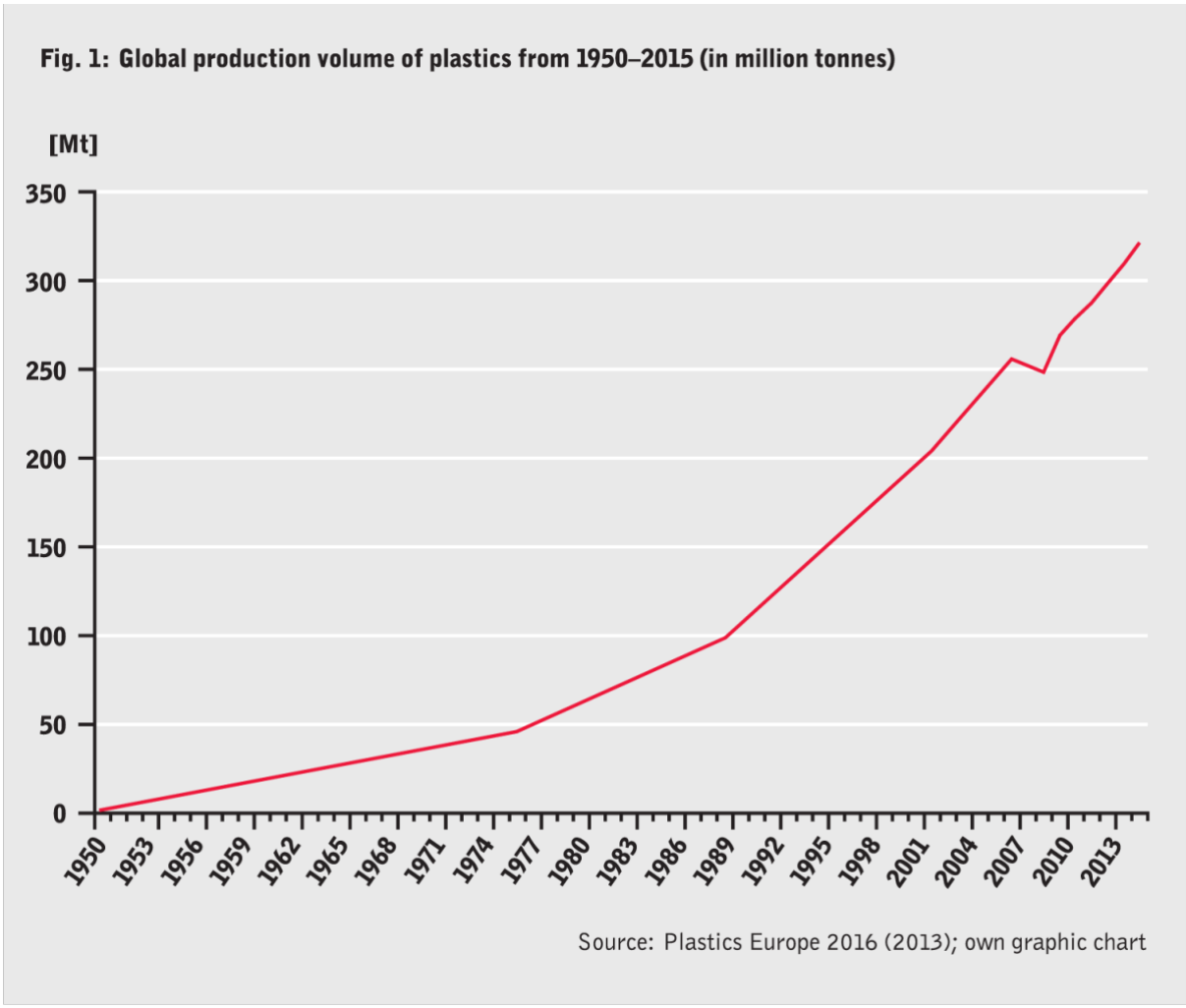

Figure 2: Global production volume of plastics 1950–2015 (in million tonnes).

Source: Simon & Schulte 2017, p. 13.

This was a period of unexpected and forced economic growth, one Alier embraced as a 'a welcome change' (p. 1099). So long as economic growth remains, companies and consumers that benefit from it (including those represented in these emerging coalitions) will resist fundamental changes to the status quo and continue to produce and consume plastic. Global economic paradigms of growth – reinforced by the powerful triumvirate of corporate media, corporate think tanks and corporate capture of political processes – represent a formidable 'system-preserving power' (Han, 2015), and thus a barrier to conventions being able to positively shape and address the problem of plastics. The only thing to have slowed this paradigm so far was a dramatic economic downturn: forced degrowth.

Larger structural barriers to global conventions, such as dominant economic paradigms of endless capitalist growth, may represent an important counterpoint to HBF's argument for a convention. Structural changes to global economic systems are difficult for a convention alone to tackle. However, the inclusion of key private sector players and bodies like the EMF is encouraging. Although not quite a degrowth model, the circular economy models espoused by the EMF are at least *partially* focused on

managing growth more sustainably. The same can be said for other groups' efforts as well. For example, sustainably managed growth can also be achieved by capturing value from waste and therefore drawing economic growth from extant (currently underutilised or wasted) resources. As stated earlier, plastic packaging alone represents some USD 80–120 billion in lost value annually (World Economic Forum, Ellen MacArthur Foundation & McKinsey & Company, 2016).

Structural challenges and looming crises related to them also present an opportunity for the kind of paradigm change Naomi Klein and others argue is necessary to achieve a sustainable global economy. The International Monetary Fund recently warned that the global economy remains critically vulnerable to another financial crisis, and that should another happen, the same level of quantitative easing and government bailouts of key financial institutions will not be possible (IMF, 2018). Alier's 'welcome change' of forced economic degrowth may return once more. The last economic crisis precipitated a dramatic rise in the interest of corporate social responsibility and sustainability more generally, as corporations struggled to increase efficiency and capture value from novel revenue streams (Benn et al., 2014). A similar crisis could potentially accelerate the creation of a convention, while simultaneously increasing the economic value of other innovations in waste recovery, recycling, and plastic alternatives. The interplay between crisis and opportunity is present here as it is within climate change: 'one of the great things about climate change is it's a great tool for bringing down the capitalist system' Naomi Klein was once paraphrased as saying (Krien, 2017, p. 89). The same may be true of plastics.

Conclusion

This paper concludes that while a convention is a critical goal, it will likely come in the medium–long term, after what will likely be some lengthy and contentious debate between stakeholders over goals, priorities and targets. Meanwhile, in the current and near-term, these same business-led innovations will continue in different directions, all of which are arguably equally important. The ACC and partners will focus on materials recovery and sustainable alternatives, whereas the EMF coalition will focus on promoting circular economy models. Other projects like Ocean Cleanup that combine philanthropy and engineering solutions will continue to offer sustainability gains in yet other areas such as environmental restoration.

A convention therefore isn't perhaps *the* next step, but something closer to a final step – one that consolidates the gains made by these groups and facilitates their successes elsewhere through capacity building measures (Pillar 3, from Figure 1). This aspect of the convention could be particularly effective in rolling out proven solutions for countries where the source of pollution is strong but the capacity to address it is weak. The creation of a plastics convention therefore is critical, but heavily complemented by equally important approaches, many of which will precede the formation of a convention, and even make such a thing possible in the first place.

Acknowledgements

With thanks to my lecturer Edwina Fingleton-Smith whose course laid the groundwork for this paper, and to fellow student Machaon Smeaton whose research helped inspire it, and to my family whose support makes university possible.

References

Alier, J M, 2009. Socially Sustainable Economic De-growth. *Development and Change,* 40(6), pp. 1099–119. doi.org/10.1111/j.1467-7660.2009.01618.x.

Benn, S, Edwards, M & Williams, T, 2014. *Organizational Change for Corporate Sustainability.* 3rd ed. Routledge: London. doi.org/10.4324/9781315819181.

Bostrom, N, 2013. Existential Risk Prevention as Global Priority. *Global Policy*, February, 4(1), pp. 15–31.

Ellen MacArthur Foundation, 2018. *Ellen MacArthur Foundation to form coalition of leaders to create circular economy for plastics.* [Online] Available at: newplasticseconomy.org/news/ellen-macarthur-foundation-to-form-coalition-of-leaders-to-create-circular-economy-for-plastics [Accessed 12 October 2018].

Ellen MacArthur Foundation & World Economic Forum, 2017. *The New Plastics Economy – Catalysing Action,* West Cowes, UK: Ellen MacArthur Foundation.

Group of Seven, 2018. *G7 Ocean Plastics Charter.* [Online] Available at: g7.gc.ca/wp-content/uploads/2018/06/OceanPlasticsCharter.pdf [Accessed 12 October 2018].

Han, B-C, 2015. *Why Revolution is No Longer Possible.* [Online] Available at: www.opendemocracy.net/transformation/byung-chul-han/why-revolution-is-no-longer-possible [Accessed 12 October 2018].

Harris, J, Brown, VA & Russel, JY, 2010. *Tackling Wicked Problems: Through the Transdisciplinary Imagination.* 1st ed. New York: Earthscan.

IMF (International Monetary Fund), 2018. *World Economic Outlook.* [Online] Available at: www.imf.org/en/Publications/WEO/Issues/2018/09/24/world-economic-outlook-october-2018 [Accessed 12 October 2018].

Jones, KR et al., 2018. The Location and Protection Status of Earth's Diminishing Marine Wilderness. *Current Biology,* 28(15), pp. 2506–12. doi.org/10.1016/j.cub.2018.06.010.

Kim, J-S, Lee, H-J, Kim, S-K & Kim, H-J, 2018. Global Pattern of Microplastics (MPs) in Commercial Food-Grade Salts: Sea Salt as an Indicator of Seawater MP Pollution. *Environmental Science and Technology*. doi.org/10.1021/acs.est.8b04180.

Kosuth, M, Mason, SA & Wattenberg, EV, 2018. Anthropogenic Contamination of Tap Water, Beer, and Sea Salt. *PLOS One,* 13(4). doi.org/10.1371/journal.pone.0194970.

Krien, A, 2017. The Long Goodbye : Coal, Coral and Australia's Climate Deadlock. *Quarterly Essay*, Issue 66.

Lebreton, L et al., 2018. Evidence that the Great Pacific Garbage Patch is Rapidly Accumulating Plastic. *Scientific Reports,* 8(4666). doi.org/10.1038/s41598-018-22939-w.

Materials Recovery for the Future, 2018. *Research Program Announces Partnership with MRF to Pilot Curb Side Recycling of Flexible Plastic Packaging.* [Online] Available at: www.materialsrecoveryforthefuture.com/press-releases/pilot-press-release/ [Accessed 12 October 2018].

National Oceanic and Atmospheric Administration, 2018. *What are Microplastics?*. [Online] Available at: oceanservice.noaa.gov/facts/microplastics.html [Accessed 12 October 2018].

Russell, S, 2017. *Toward a Broader View of the New Plastics Economy*. [Online] Available at: www.plasticstoday.com/packaging/toward-broader-view-new-plastics-economy/211530339956884 [Accessed 12 October 2018].

Simon, N & Schulte, M, 2017. *Stopping Global Plastic Pollution: The Case for an International Convention*, Germany: Heinrich Böll Foundation.

The Ocean Cleanup, 2018. *The Impact of Cleanup.* [Online] Available at: www.theoceancleanup.com/ [Accessed 12 October 2018].

Trucost, 2016. *Plastics and Sustainability: A Valuation of Environmental Benefits, Costs and Opportunities for Continuous Improvement*. USA: Trucost.

US Geological Survey, 2016. *How Much Water is There on, in, and above the Earth?*. [Online] Available at: water.usgs.gov/edu/earthhowmuch.html [Accessed 12 October 2018].

Veiga, J et al., 2016. *Identifying Sources of Marine Litter: MSFD GES TG Marine Litter Thematic Report*. Luxembourg: JRC Technical Report; EUR 28309.

World Economic Forum, Ellen MacArthur Foundation & McKinsey & Company, 2016. *The New Plastics Economy – Rethinking the Future of Plastics*. [Online] Available at: www.ellenmacarthurfoundation.org/publications/the-new-plastics-economy-rethinking-the-future-of-plastics [Accessed 12 October 2018].

Trump's trade tariffs and the Article XXI 'trump' card

GEORGIE JUSZCYK

Abstract

The recent 'Trump tariffs' have heralded concerns about a blooming 'trade war' between the US and China. These unilateral measures are a threat to the substance of international trade law. In part, this is due to questions the measures raise about the interpretation of the Article XXI 'Security Exception' in the General Agreement on Tariffs and Trade. It is feared that an expansive reading of Article XXI would undermine the World Trade Organization (WTO) by allowing 'anything under the sun' to be considered 'necessary' for the protection of 'essential security interests'.[1] This paper argues that WTO Members should respond by narrowing the scope of the exception. This paper also acknowledges, however, that too narrow a reading of the exception also poses problems, as this would reduce the ability of member states to introduce legitimate regulations in regard to national security interests.

Introduction

US President Trump's tariffs impose rates of 25 per cent and 10 per cent on steel and aluminium imports from Norway, Mexico, Canada, the EU, India and China.[2] These unilateral tariffs are a threat to the substance of international trade law. Among other reasons, this is because the Trump administration has invoked the Article XXI security exceptions in defence of complaints by the affected countries.[3] It is feared that an expansive reading of Article XXI would undermine the World Trade Organization

---

[1] Roger P Alford, 'The Self-Judging WTO Security Exception' (2011) 3 *Utah Law Review* 697, 698; *Summary of the Twenty-Second Meeting*, CP.3/SR22-II/28 (8 June 1949) (Summary of the Record of the Twenty-Second Meeting, Third Session); Johannes Fahner, 'Qatar under Siege: Chances for an Article XXI Case?' *EJIL: Talk!* (online) 9 January 2018 <www.ejiltalk.org/qatar-under-siege-chances-for-an-article-xxi-case/>; Raj Bhala, 'National Security and International Trade Law: What the GATT Says, And What the United States Does' (1998) 19(2) *National Security and International Trade Law*, 275; Peter Lindsay, 'The Ambiguity of GATT Article XXI: Subtle Success or Rampant Failure?' (2003) 52 *Duke Law Journal*, 1294; Ji Yeong Yoo and Dukgeun Ahn, 'Security Exceptions in the WTO System: Bridge or Bottle-Neck for Trade and Security?' (2016) 19 *Journal of International Economic Law* 423.

[2] 'A summary of US President Donald Trump's many trade wars', *Strait Times* (online) 6 July 2018 <www.straitstimes.com/world/united-states/a-summary-of-us-president-donald-trumps-many-trade-wars>; Simon Lester, 'Litigating GATT Article XXI: The US View of the Scope of the Exception' on *International Economic Law and Policy Blog* (19 March 2018) <worldtradelaw.typepad.com/ielpblog/2018/03/litigating-gatt-article-xxi-the-us-view-of-the-scope-of-the-exception.html>.

[3] 'A summary of US President Donald Trump's many trade wars', above n 2.

(WTO) by allowing 'anything under the sun' to be considered 'necessary' for the protection of 'essential security interests'.[4] This would mean that Trump's restrictive trade measures – which would otherwise be in breach of Articles I, II, X, XI and XIX of the General Agreement on Trade and Tariffs (GATT) – would get a free pass, defeating the WTO's objective of promoting trade liberalisation and its associated socio-economic benefits.[5] This paper argues that WTO Members should respond by narrowing the scope of the exception. To demonstrate this point, this paper will first explain how Article XXI is currently understood and why it needs to be changed, before exploring the different methods of doing so.

How is Article XXI currently understood?

Currently, Article XXI is regarded as a self-judging provision.[6] Consequently, measures enacted under this exception are seen as beyond the jurisdiction of the WTO's Dispute Settlement Body.[7] Even those who argue that Article XXI is technically reviewable despite its self-judging nature – distinguishing between the 'authority to define' and the 'authority to interpret' – acknowledge that such a right holds 'no practical importance'.[8]

This conclusion is largely derived from interpretations of Article XXI's chapeau and part (b). The chapeau reads, '[n]othing in this agreement shall be construed …'[9] When drafted, this provision included words of limitation preventing measures that were an 'arbitrary or a disguised restriction on international trade.'[10] Today, that language is located in the Article XX's chapeau but is conspicuously absent from Article XXI.[11] This suggests that the limitations no longer apply, lending support to Article XXI's 'all-encompassing' and self-judging nature.[12] This interpretation is further strengthened by the lack of an explicit definition of 'essential security interests' in part (b).[13] The interpretative difficulties surrounding parts (a), (c) and the subparagraphs of part (b) will not be explored for the

---

[4] Alford, above n 1, 698; *Summary of the Twenty-Second Meeting*, CP.3/SR22-II/28 (8 June 1949) (Summary of the Record of the Twenty-Second Meeting, Third Session); Fahner, above n 1; Bhala, above n 1, 275; Lindsay, above n 1, 1294; Yoo and Ahn, above n 1, 423.

[5] *Summary of the Twenty-Second Meeting*, CP.3/SR22-II/28 (8 June 1949) (Summary of the Record of the Twenty-Second Meeting, Third Session); Alford, above n 1, 698; Fahner, above n 1; Bhala, above n 1, 275; Lindsay, above n 1, 1294; Yoo and Ahn, above n 1, 423.

[6] Alford, above n 1, 704; Bhala, above n 1, 269.

[7] Lindsay, above n 1, 1291; Yoo and Ahn, above n 1, 427.

[8] Bhala, above n 1, 267; Yoo and Ahn, above n 1, 428.

[9] *General Agreement on Tariffs and Trade 1994*, opened for signature 30 October 1947, 55 UNTS 187 (entered into force 1 January 1948) ('*GATT*'), Article XXI.

[10] Yoo and Ahn, above n 1, 421.

[11] *GATT*, Article XX; Yoo and Ahn, above n 1, 421.

[12] Fahner, above n 1; Yoo and Ahn, above n 1, 422; Bhala, above n 1, 268; Lindsay, above n 1, 1291.

[13] Yoo and Ahn, above n 1, 423; Lindsay, above n 1, 1278.

purpose of this essay, as the heart of the controversy around Article XXI is how best to limit the scope of a 'self-judging' exception.

Article XXI's and GATT's interests are also diametrically opposed. There is no implied definition as per the 'ordinary meaning' of the words, against the background of the 'context, object and purpose' of Article XXI and the GATT.[14] Rather, 'essential security interests' are to be defined by the Member themselves, as indicated by the words 'it [the nation] considers' (its subjective opinion).[15] In contrast to Article XXI, such subjective language is absent from Article XX, which instead uses objective language – for example, 'necessary to protect human, animal or plant life or health.'[16] Additionally, in *China – Raw Materials* it was found that Article XI:2(a) was *not* self-judging because otherwise, 'Article XI:2 could have been drafted in a way such as Article XXI(b) …'[17]

Given the importance of state practice to the WTO, it is also useful to turn to it for confirmation.[18] The majority of states strongly affirm Article XXI's self-judging nature, citing the need to prioritise security over trade interests.[19] In 1949, Czechoslovakia's suggestion for a narrow reading of Article XXI was rejected in favour of an approach which championed the rights of each country to make their own determinations on what constituted a security interest under the exception,[20] while in 1961, Ghana used a similar rationale to defend a boycott on Portuguese goods, asserting its right to be the 'sole judge' of its own security.[21] In *Trade Restrictions Affecting Argentina Applied for Non-Economic Reasons* ('*Argentina*'), the EU, Australia and Canada asserted their 'inherent rights, of which Article XXI of the General Agreement was a reflection' and reiterated that Article XXI was beyond the GATT Panel's purview.[22] This was met by approbation from 20 of the 37 other countries (11 did not discuss the issue and 6 disagreed).[23] Finally, in *United States – Trade Measures Affecting Nicaragua* ('*Nicaragua*') the US's defence of its ban on imports of Nicaraguan origin under Article XXI was approved by 19 of 43

---

[14] *Vienna Convention on the Law of Treaties*, opened for signature 23 May 1969, 1155 UNTS 331 (entered into force 27 January 1980), Article 31(1).

[15] Fahner, above n 1; Matthew Kahn, 'Pretextual Protectionism? The Perils of Invoking the WTO National Security Exception' on *Lawfare* (21 July 2018) <www.lawfareblog.com/pretextual-protectionism-perils-invoking-wto-national-security-exception>; Yoo and Ahn, above n 1, 427; Alford, above n 1, 698; Lindsay, above n 1, 1282; Bhala, above n 1, 268; Lester, above n 2.

[16] Lindsay, above n 1, 1282.

[17] Panel Report, *China – Measures Related To The Exportation Of Various Raw Materials*, WT/DS394/R (5 July 2011), 7.276.

[18] *Vienna Convention on the Law of Treaties*, opened for signature 23 May 1969, 1155 UNTS 331 (entered into force 27 January 1980), Article 31(3)(b); Lindsay, above n 1, 1282; Alford, above n 1, 706; Joost Pauwelyn, 'The Role of Public International Law in the WTO: How Far Can We Go? (2001) 95(3) *The American Journal of International Law*, 536.

[19] Bhala, above n 1, 317; Yoo and Ahn, above n 1, 419; Alford, above n 1, 708.

[20] Alford, above n 1, 709; Bhala, above n 1, 268; Lindsay, above n 1, 1286.

[21] Kahn, above n 15; Lindsay, above n 1, 1286; Bhala, above n 1, 269; Alford, above n 1, 732.

[22] Panel Report, *Trade Restrictions Affecting Argentina Applied for Non-Economic Reasons*, WTO Doc L/5317 (30 April 1982).

[23] Alford, above n 1, 710.

nations (15 did not discuss the issue and 9 disagreed).[24] The collective weight of this state practice is compelling support for a self-judging interpretation of the security exception.

Finally, the persuasive judgements of other courts, such as the International Court of Justice (ICJ), also support Article XXI's 'self-judging' interpretation. Most notably in *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America)* the ICJ found that it had jurisdiction to interpret a security exception in a treaty between the disputing parties. This decision was reached by distinguishing the treaty provision from Article XXI. Instead of 'it considers necessary', the treaty allowed measures that were 'necessary', indicating that the treaty provision was *not* self-judging.[25] Again, turning to such bodies is useful because the WTO's dispute resolution mechanisms exist within the broader context of international law.[26]

Why does Article XXI need to be changed?

One school of thought about why Article XXI needs to be amended concerns pro–status quo preferences to rely on other means of rule. It is argued that the 'sovereignty safety valves' already built into the WTO system adequately provide for Members' needs to further their security objectives without turning to Article XXI.[27] These 'valves' include: procedures established by the *Decision Concerning Article XXI of the General Agreement,* such as the need to give a priori notice of an impending measure; the opt-out option in accession rules, which can be used against unfriendly nations; preferential trading agreements, which can be used for friendly nations; 'strings attached' preferential agreements between developed and developing countries and, finally, non-violation claims under Article XXIII.[28] The latter has been a particular focus of the literature surrounding the exception, although it is now largely settled that the right will apply even in the context of Article XXI.[29] Members may also be circumspect in their invocation of Article XXI because they enjoy the flexibility and control afforded by the provision's 'constructive ambiguity'.[30] In order to prevent this discretion from being stripped away, they are careful

---

[24] Panel Report, *United States – Trade Measures Affecting Nicaragua*, WTO Doc L/6053 (13 October 1986); Kahn, above n 15; Michael Woods, 'GATT Article XXI's National Security Exception – The Ultimate Trade Policy Conundrum' *Woods Lafortune LLP* (online), 9 March 2018 <www.wl-tradelaw.com/gatt-article-xxis-national-security-exception-the-ultimate-trade-policy-conundrum/>; Lindsay, above n 1, 1285; Simon Lester, 'Non-Violation Claims in the Steel/Aluminium WTO Complaints' on *International Economic Law and Policy Blog* (20 June 2018) <worldtradelaw.typepad.com/ielpblog/2018/06/the-steelaluminum-wto-complaints.html>; Lester, above n 2; Alford, above n 1, 713.

[25] Fahner, above n 1; Alford, above n 1, 707; Yoo and Ahn, above n 1, 427; Lindsay, above n 1, 1285.

[26] Pauwelyn, above n 18, 577; Lindsay, above n 1, 1280.

[27] Alford, above n 1, 750.

[28] *Decision Concerning Article XXI of the General Agreement* L/5246 (2 December 1982) (Text of Members' Decision), Article 1; Alford, above n 1, 725; Fahner, above n 1.

[29] Lindsay, above n 1, 1280; Bhala, above n 1, 278; Alford, above n 1, 746.

[30] Lindsay, above n 1, 1279; Bhala, above n 1, 272; Fahner, above n 1.

in their invocation and instead prefer resolving issues through 'international pressure and diplomacy'.[31] Indeed, the exception has not often been raised or used recklessly, despite the decided lack of other coercive, rational or normative motives in preventing Members from doing so.[32] However, the very fact that the Trump tariffs now exist seems to discount this argument.

Perhaps we are now in a world where Article XXI is no longer tolerated because parties no longer share 'mostly … common security goals'.[33] Additionally, although the notice procedures and preferential trading agreements are now common practice, other mitigating factors do not apply as well. The opt-out measures can only be applied during accession and so are useless in situations where relations sour between countries that are already Members. This is the more likely situation, given the WTO now has 153 Members out of the world's 195 countries.[34] Meanwhile, non-violation claims are rarely invoked (then again, Article XXI invocations are almost as rare). In the case of the Trump tariffs, only Mexico and India included a non-violation claim under Article XXIII:1(b).[35]

While it is clear that change to Article XXI needs to occur, it is unclear in which direction this change needs to happen, given the situation of minority state powers. Too broad a reading would invite the normalisation and legitimisation of the exception's abuse by allowing 'anything under the sun' – an allowance particularly likely to be exploited by economically powerful states.[36] This has long been a concern of the minority in state practice – in the *Argentina* case, Argentina named Article XXI as a 'magnificent safeguard clause', saying '[i]t would appear that trade restrictions could be adopted without having to be justified or approved', while Brazil warned against setting a 'dangerous precedent' by not requiring a concrete demonstration of 'essential security interests'.[37] In *Nicaragua,* a handful of Members (including Poland, Cuba and Nicaragua itself) argued that the US's (self-)judgement that Nicaragua posed a threat to its national security was ridiculous – a global superpower could not be threatened by such a relatively small and underdeveloped nation.[38]

Too narrow a reading, however, would also undermine the WTO regime. Article XXI was an opt-out provision designed to increase buy-in to the WTO by allowing Members autonomy over 'political

---

[31] Lindsay, above n 1, 1279; Bhala, above n 1, 272; Fahner, above n 1.

[32] Alford, above n 1, 750; Bhala, above n 1, 271; Yoo and Ahn, above n 1, 439.

[33] Yoo and Ahn, above n 1, 433.

[34] Bhala, above n 1, 271.

[35] Lester, above n 2.

[36] *Summary of the Twenty-Second Meeting*, CP.3/SR22-II/28 (8 June 1949) (Summary of the Record of the Twenty-Second Meeting, Third Session); Alford, above n 1, 698; Fahner, above n 1; Bhala, above n 1, 275; Lindsay, above n 1, 1294; Yoo and Ahn, above n 1, 423.

[37] Panel Report, *Trade Restrictions Affecting Argentina Applied for Non-Economic Reasons*, WTO Doc L/5317 (30 April 1982).

[38] Panel Report, *United States – Trade Measures Affecting Nicaragua,* WTO Doc L/6053 (13 October 1986).

matters' seen to be beyond the gamut of the WTO's institutional competence.[39] Indeed, in *Argentina* the US stated that the self-judging approach was more likely to preserve the longevity of the WTO system, since no country would submit to an organisation which might decide against its national interests.[40] However, the WTO today is not the same 'pseudo-UN agency' envisioned in the design of the International Trade Organization (ITO) – on the contrary, it presides over an increasing number of progressively complex trade issues.[41] Hence, this paper iterates that it is well within the WTO's capability to have a hand in limiting the scope of the Article XXI exception. Further, although there are a number of ways in which this Article XXI 'loophole' can be closed, it must be done with a view towards 'balance', as per the original intentions of the drafters.[42]

How should Article XXI be understood?

Textual solutions

Good faith

One method of limiting Article XXI's scope is to imply a 'good faith' requirement that is subject to WTO review.[43] Some go further, suggesting that Article XXI should be amended to explicitly include this limitation.[44] The principle of good faith is supported by the GATT's text, whose 'context, object and purpose' advocates for trade restrictions on limited grounds with proportionate applications.[45] It is also supported by state practice given that, until now, the exception has largely been applied in good faith, with a mere 'margin of discretion' – whether this be due to a fear of reprobation, self-interest or other normative values it is not known.[46] Statements such as the EU's oral submission in *Russia – Traffic in Transit,* which notes that any review should be limited to questions of good faith, further strengthen this argument.[47]

The principle of good faith has also been applied in comparable bodies of international law. It has been recognised in *Gabcikovo-Nagymoros* and applied specifically in relation to a 'self-judging' security

---

[39] Fahner, above n 1;Woods, above n 24; Lindsay, above n 1, 1279.

[40] Panel Report, *Trade Restrictions Affecting Argentina Applied for Non-Economic Reasons*, WTO Doc L/5317 (30 April 1982); Alford, above n 1, 712; Lindsay, above n 1, 1286; Bhala, above n 1, 269.

[41] Yoo and Ahn, above n 1, 440; Alford, above n 1, 702.

[42] Kahn, above n 15; Alford above n 1, 699; Yoo and Ahn, above n 1, 426; Bhala, above n 1, 269; Fahner, above n 1.

[43] Lester, above n 2; Fahner, above n 1;Lindsay, above n 1, 1307; Yoo and Ahn, above n 1, 429; Alford, above n 1, 705.

[44] Yoo and Ahn, above n 1, 429.

[45] *Vienna Convention on the Law of Treaties*, opened for signature 23 May 1969, 1155 UNTS 331 (entered into force 27 January 1980), Article 31(1); Alford, above n 1, 706.

[46] Alford, above n 1, 708; Woods, above n 24; Bhala, above n 1, 272

[47] European Union Third Party Written Submission, *Russia – Measures Concerning Traffic in Transit*, WTO Doc WT/DS512/Ares(2017)5434182 (8 November 2017).

exception in *Djibouti v France*.[48] The principle has also been crystallised by its inclusion in Article 31 of the Vienna Convention on the Law of Treaties.[49] When applying this to Trump's tariffs, however, its use would likely be limited. Article XXI(a)'s protection of the non-disclosure of sensitive information would likely operate to both obstruct any dispute settlement process and make a lack of good faith difficult to prove, while any test for the principle remains difficult to define.[50]

Objective–subjective

The objective–subjective approach suggests an alternative interpretation of Article XXI that fuses objective and subjective standards. Herein, Members 'consider' whether the measure is 'necessary' but the reasonableness of the measures are subject to review.[51] This proportionality test asks whether a 'reasonable, similarly-situated government … faced with the same circumstances' would invoke the exception.[52] It could include considerations such as: whether or not the measure is effective (although national security sanctions are generally regarded to be 'ineffectual at best, and counter-productive at worst'[53]); how effective the measure is in proportion to the burden it imposes on a third party; whether that third party is a developing country and how consistently the measure applies to other nations in like circumstances.[54] This approach satisfies the need for objective criteria while also accommodating the self-judging language of Article XXI (although there is some debate about whether this quasi-accommodation will suffice).[55]

The objective–subjective approach also offers more concrete criterion than the good faith principle, despite not being such a widely accepted idea. However, some support for this approach can be found in state practice, under items such as the *Decision Concerning Art XXI of the General Agreement*, which urges Members to take into account the interests of the third parties and balance security interests with disruption to international trade.[56] Another example is the language of the GATT Panel in *Nicaragua,* that it is incumbent on 'each contracting party, whenever it made use of its rights under Article XXI, [to] carefully weigh its security needs against the need to maintain stable trade relations'.[57] Additionally, scholars have argued that if the final decision lay with the party invoking Article XXI, it would be

---

[48] Fahner, above n 1.

[49] Yoo and Ahn, above n 1, 442.

[50] *GATT*, Article XXI(a); Kahn, above n 15.

[51] Alford, above n 1, 704; Yoo and Ahn, above n 1, 442; Lindsay, above n 1, 1286.

[52] Bhala, above n 1, 275.

[53] Ibid, 266.

[54] Fahner, above n 1; Alford, above n 1, 706; Lindsay, above n 1, 1307; Yoo and Ahn, above n 1, 428; Bhala, above n 1, 276.

[55] Yoo and Ahn, above n 1, 429; Fahner, above n 1.

[56] *Decision Concerning Article XXI of the General Agreement*, L/5246 (2 December 1982) (Text of Members' Decision).

[57] Panel Report, *United States – Trade Measures Affecting Nicaragua*, WTO Doc L/6053 (13 October 1986); Alford, above n 1, 716; Woods, above n 24; Kahn, above n 15.

ineffective in creating any legal limitation on a country's power to escape sanction, as originally intended.[58] Otherwise, they say, if the provision was completely self-judging, the words of limitation contained in the subparagraphs of Article XXI(b) would be rendered useless.[59]

Amend the chapeau

Alternatively, it has been suggested that the chapeau of Article XXI could be amended in order to impose a 'trade restrictiveness' condition, similar to that in Article XX.[60] However, as discussed above, the removal of this test during the original drafting process reflects a decision by countries that the exception would otherwise be too narrow to effectively protect countries' 'essential security interests'. Therefore, this amendment would unlikely be supported.[61]

Non-textual solutions

UN/WTO collaboration

A more novel albeit unsupported idea suggests procedural restraints, based on UN–WTO collaboration. One version of this establishes a WTO – Security Council Committee that could offer non-binding, non-precedential opinions or advice on counter-retaliatory measures.[62] Another version is to encourage limiting Article XXI to scenarios where a UN Security Council resolution highlighting the relevant security interest has already been issued.[63] Both of these solutions aim to resolve the issue of the WTO's alleged lack of institutional competence over 'political matters'; however, both will likely be unpopular. Although non-binding, the history of Members' obedience to WTO recommendations indicates that these statements would still be persuasive and ignoring them would still be detrimental to WTO credibility. Besides, the US already complains of WTO overreach.[64]

Conclusion

This paper has argued that, the current interpretation of Article XXI is too broad. A more balanced, narrower reading of Article XXI – through either the 'good faith' or 'objective–subjective' solutions – is needed in order to protect the substance and future of international trade law. Contrary to claims that the Article's 'constructive ambiguity' is actually beneficial, it is clear that the current understanding of

---

[58] Alford, above n 1, 706.

[59] Lindsay, above n 1, 1287.

[60] *GATT,* Article XX; Alford, above n 1, 705; Yoo and Ahn, above n 1, 442.

[61] Woods, above n 24.

[62] Bhala, above n 1, 276.

[63] Ibid, 267.

[64] Jack Ewing, 'Europe Feels the Squeeze of the Trump Trade Tariffs', *New York Times* (online) 2 April 2018 <www.nytimes.com/2018/08/02/business/economy/europe-trade-trump-tariffs.html>.

Article XXI – as a self-judging provision with no capacity for review under the WTO's dispute resolution system – will not hold. The issues raised by Trump's tariffs are not new. It is merely that the long-held consensus on careful invocations of Article XXI has been undone by Trump's tariffs. While it is possible that future American administrations may reverse Trump's tactics, reducing the urgency of such a question, the fact is that the precedent has been made. If not answered soon, the question of whether a country will next try to invoke the security exceptions is more a matter of 'when' and not 'if'.

Bibliography

Articles/books/reports

Alford, Roger P, 'The Self-Judging WTO Security Exception' (2011) 3 *Utah Law Review* 697

Bhala, Raj, 'National Security and International Trade Law: What the GATT Says, And What the United States Does' (1998) 19(2) *National Security and International Trade Law* 263

Lindsay, Peter, 'The Ambiguity of GATT Article XXI: Subtle Success or Rampant Failure?' (2003) 52 *Duke Law Journal* 1277

Pauwelyn, Joost, 'The Role of Public International Law in the WTO: How Far Can We Go? (2001) 95(3) *The American Journal of International Law* 535

World Trade Organization, *WTO Analytical Index: Guide to WTO Law and Practice* (World Trade Law Organisation, 3rd ed, 2011)

Yoo, Ji Yeong, and Dukgeun Ahn, 'Security Exceptions in the WTO System: Bridge or Bottle-Neck for Trade and Security?' (2016) 19 *Journal of International Economic Law* 417

Cases

European Union Third Party Written Submission, *Russia – Measures Concerning Traffic in Transit* WTO Doc WT/DS512/Ares(2017)5434182 (8 November 2017)

Panel Report, *China – Measures Related To The Exportation Of Various Raw Materials*, WTO Doc WT/DS394/R (5 July 2011).

Panel Report, *Trade Restrictions Affecting Argentina Applied for Non-Economic Reasons*, WTO Doc L/5317 (30 April 1982).

Panel Report, *United States – Trade Measures Affecting Nicaragua* WTO Doc L/6053 (13 October 1986).

Treaties

General Agreement on Tariffs and Trade 1994, opened for signature 30 October 1947, 55 UNTS 187 (entered into force 1 January 1948) (*'GATT'*)

Vienna Convention on the Law of Treaties, opened for signature 23 May 1969, 1155 UNTS 331 (entered into force 27 January 1980)

Other

'A Summary of US President Donald Trump's Many Trade Wars', *The Straits Times* (online) 6 July 2018 <www.straitstimes.com/world/united-states/a-summary-of-us-president-donald-trumps-many-trade-wars>

Decision Concerning Article XXI of the General Agreement, L/5246 (2 December 1982) (Text of Members' Decision)

Donnan, Shawn, 'Trump's Hopes for EU Trade Deal Hampered by Agriculture', *Financial Times* (online), 3 August 2018, <www.ft.com/content/dfe8a8a6-95db-11e8-b67b-b8205561c3fe>

Euronews, 'Juncker Responds to Trump's Trade Tariffs: "We Can Also Do Stupid"', *Euronews* (online), 3 March 2018 <www.euronews.com/2018/03/03/juncker-responds-to-trump-s-trade-tariffs-we-can-also-do-stupid->

Ewing, Jack, 'Europe Feels the Squeeze of the Trump Trade Tariffs', *New York Times* (online) 2 April 2018 <www.nytimes.com/2018/08/02/business/economy/europe-trade-trump-tariffs.html>

Fahner, Johannes, 'Qatar Under Siege: Chances for an Article XXI Case?' *EJIL: Talk!* (online) 9 January 2018 <www.ejiltalk.org/qatar-under-siege-chances-for-an-article-xxi-case/>

Kahn, Matthew, 'Pretextual Protectionism? The Perils of Invoking the WTO National Security Exception', on *Lawfare* (21 July 2018) <www.lawfareblog.com/pretextual-protectionism-perils-invoking-wto-national-security-exception>

Lester, Simon, 'Litigating GATT Article XXI: The US View of the Scope of the Exception' on *International Economic Law and Policy Blog* (19 March 2018) <worldtradelaw.typepad.com/ielpblog/2018/03/litigating-gatt-article-xxi-the-us-view-of-the-scope-of-the-exception.html>

Lester, Simon, 'Non-Violation Claims in the Steel/Aluminium WTO Complaints' on *International Economic Law and Policy Blog* (20 June 2018) <worldtradelaw.typepad.com/ielpblog/2018/06/the-steelaluminum-wto-complaints.html>

Palmer, Doug, 'US Sides with Russia in WTO National Security Case Against Ukraine', *Politico* (online), 30 July 2018 <www.politico.eu/article/us-sides-with-russia-in-wto-national-security-case-against-ukraine/>

Summary of the Twenty-Second Meeting, CP.3/SR22-II/28 (8 June 1949) (Summary of the Record of the Twenty-Second Meeting, Third Session)

Woods, Michael, 'GATT Article XXI's National Security Exception – The Ultimate Trade Policy Conundrum', *Woods Lafortune LLP* (online), 9 March 2018, <www.wl-tradelaw.com/gatt-article-xxis-national-security-exception-the-ultimate-trade-policy-conundrum/>

Grappling with dissent and political silencing during transition: The consequences of Myanmar's NLD-led government defamation law

JAMES ATKINSON

Abstract

In 2013, Myanmar's government passed the Telecommunications Law to regulate private telecommunication providers. Article 66(d) was intended to provide protection against defamation. Under the National League for Democracy (NLD)-led government, however, this Article is being used to stifle political dissent and suppress freedoms of media and expression in digital spaces. Wide-ranging public campaigns led by media and civil actors have advocated for the repealing of this Article, though none have been successful. This paper considers the barriers to reforming Article 66(d). It draws on primary interviews with activists, media professionals and academics based in Myanmar alongside secondary research to identify three key barriers to reform: Myanmar's political context under an NLD-led government, cultural dynamics towards the media, and existing policy frameworks. While this analysis finds that there are potential avenues for reforming the Article, it makes the conclusion that it is unlikely that meaningful reform will take place in the current context. This research provides an interesting insight to how Southeast Asian countries like Myanmar reconcile their political transitions and development with the influence of Western and domestic media.

Introduction

The National League for Democracy's (NLD) win at Myanmar's 2015 general election came with the expectation that it would further liberalise freedoms of media and expression. Under the NLD's governance, however, these freedoms have regressed.[1] A symbol of this decline is the use of Article 66(d) of the 2013 Telecommunications Law. The Telecommunications Law regulates private telecommunication providers and Article 66(d) contains provisions to charge an individual for using a

[1] Free Expression Myanmar, 'Myanmar's media freedom at risk' (Report, Yangon, Myanmar, 2018) 9–11.

provider in a defamatory way.[2] The intention was to provide respite for a country grappling with the influence of a liberalising media.

The practical use of Article 66(d) tells a different story. It is used to charge individuals who communicate political dissent via social media. Since the NLD came into power, more than half of the charges under the Article have been taken against those who criticise government officials.[3] The Article continues to be used for these purposes despite significant campaigning from civil society actors who voice their concerns that it is being used as a tool to stifle freedoms of media and expression. The government displays an unwillingness for meaningful reform.[4]

This paper considers the possibility for reform under Article 66(d) by assessing three key barriers: Myanmar's political context; the country's cultural dynamics; and the impact of current policy frameworks. First, the effectiveness of previous amendments to the Article will be assessed, before drawing on current primary and secondary research to assess the possibility for reform. Ultimately, this paper argues that, without significant change in these areas, Article 66(d) of the Telecommunications Law is unlikely to result in meaningful reform.

Myanmar's political context

While there are some indications that reform may be possible in the current Myanmar parliament, actions suggests otherwise. The government amended Article 66(d) in 2017 following public outcry but the reforms were superficial. Rather than adhering to calls from civil campaigners and changing the Article's references to defamation, the most significant change was a reduction of the maximum prison sentence from three to two years.[5] These amendments indicate the government's willingness to allow public pressure to influence policy and decision-making.[6] Yet it is unlikely to make further changes in its current term.

Analysts argue that reforming Article 66(d) again is not a government priority.[7] Media and political experts in Myanmar posit that the NLD uses its parliamentary majority to delegitimise policy debates

---

[2] The Telecommunications Law 2013 (Myanmar).

[3] 'Home page', SayNoto66d, last modified 8 March 2018, accessed 28 July 2018, www.saynoto66d.info/.

[4] Tea Circle Oxford, 'Myanmar's freedom of expression as a broken promise of the NLD', 30 May 2018 teacircleoxford.com/2018/05/30/myanmars-freedom-of-expression-as-broken-promise-of-nld/.

[5] Oliver Spencer, Zar Chi Oo and Yin Yadanar Thein, 'The 66(d) amendment: Tinkering at the edges', *Frontier Myanmar*, 15 September 2017, frontiermyanmar.net/en/the-66d-amendment-tinkering-at-the-edges.

[6] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[7] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018 and Anonymous discussion with political researcher during ANU course dinner in Naypyidaw, Myanmar, 10 July 2018.

that contradict its platform. The government's intentional delay of amendments to the Peaceful Assembly and Peaceful Procession Law highlight its unilateral power to determine any law desired, wherein reforming Article 66(d) is undesirable.[8] Some NLD politicians even argued in parliament that the 2017 amendments to the Article were insignificant, giving hope to those campaigning for change.[9] Although there are some avenues for achieving reform in the current parliament, it is unlikely unless those working in and outside the NLD can change the party's platform. The NLD could respond to public pressure and meaningfully reform 66(d), yet their pattern of decision-making indicate that it is unlikely.

Others attribute the NLD's ascension to power as context for its unwillingness to reform Article 66(d). When the Union Solidarity and Development Party (USDP) won the 2010 election it did not have significant levels of public support, so it passed a suite of reforms to relax freedoms of media and expression.[10] The intention was to legitimise itself in the eyes of the public and the media by engaging them in its governance. On the contrary, the NLD interpreted its victory in 2015 as proof of its legitimacy and has demonstrated limited interaction with the media or civil society.[11] This position connects strongly with the use of Article 66(d) and the government's subsequent unwillingness to reform because they view critics as distorting their relationship with the public.[12] While activists and media professionals argue that the NLD's relationship with voters – and Myanmar's democratic transition – would strengthen if it opened itself to criticism, they agree it is unlikely given its recent electoral positioning.[13]

The NLD's position is not unique to Myanmar, rather, it is indicative of a broader trend across Southeast Asia. The influence of social media has led governments in the region to pass laws that restrict digital

---

[8] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[9] Ei Cherry Aung, 'Myanmar's parliament debates controversial telecom law', *Mizzima*, 29 July 2017, www.mizzima.com/news-features/myanmar%25E2%2580%2599s-parliament-debates-controversial-telecom-law.

[10] Nicholas Farrelly and Chit Win, 'Inside Myanmar's turbulent transformation', *Asia & the Pacific Policy Studies* 3, no. 1 (2016): 43.

[11] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; Anonymous interview with peace process consultant. Interview by ANU students. In-person interview. Yangon, Myanmar, 18 July 2018; and Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[12] Maung Aung Myoe, 'The NLD and Myanmar's foreign policy', *Journal of Current Southeast Asian Affairs* 36, no. 1 (2017): 117; Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; and Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[13] Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018 and Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

freedoms.[14] Thailand recently passed laws which threaten to prosecute social media users for criticising the monarchy – where charges are made under similar guises of 'defamation'.[15] The Vietnamese government has gone further to work with Facebook to explore strategies on the site itself that would complement the government's censorship laws.[16] These governments use morality to justify their decisions and argue that they are protecting their constituencies, but the actions taken are being used to stifle dissent.[17] This demonstrates how governments are grappling to reconcile their political transitions with Western influences and domestic criticism via social media. While the NLD may face domestic pressure to reform Article 66(d), reform may be unlikely without a broader push to relax restrictions that exist elsewhere in the region.

Even in Myanmar, Article 66(d) does not exist in isolation – it is a key element to the NLD's broader pattern of stifling dissent. Since coming into power, the NLD has made concerted efforts to silence criticism of its government on and offline. Maung Saungkha, an activist jailed under Article 66(d), notes an increasing government surveillance over media and activist circles under the NLD.[18] He says that, where the Article polices criticism in the media, the influence of the Peaceful Assembly and Peaceful Procession Law imitates similar practices in offline spaces. These laws are creating a culture of unease and self-censorship among civil actors when they engage with, or criticise, domestic politics. These feelings of uncertainty are increased by the government's recent announcement to spend six billion kyat on a social media monitoring team: while it says the team will safeguard young users, little has been said publicly about its mandate.[19] The limited transparency is worrisome to civil actors because they

[14] Liu Yangyue, 'Transgressiveness, civil society and internet control in Southeast Asia', *The Pacific Review* 27, no. 3 (2014): 384 and Ben Doherty, 'Silence of the dissenters', *The Guardian*, 22 October 2010, www.theguardian.com/technology/2010/oct/21/internet-web-censorship-asia.

[15] Yangyue, 'Transgressiveness, civil society and internet control in Southeast Asia', 384 and 'Freedom on the net 2016: Thailand country profile', Freedom House, accessed 28 July 2018, freedomhouse.org/report/freedom-net/2016/Thailand.

[16] Geoffrey Cain, 'Kill one to warn one hundred', *The International Journal of Press/Politics* 19, no. 1 (2013): 96 and David Hutt, 'Is Facebook helping Vietnam suppress online dissent?', *Asia Times*, 2 June 2018, www.atimes.com/article/is-facebook-helping-vietnams-pro-democracy-crackdown/ (site discontinued).

[17] Yangyue, 'Transgressiveness, civil society and internet control in Southeast Asia', 398.

[18] Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018 and Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[19] Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; Tea Circle Oxford, 'Myanmar's freedom of expression as a broken promise of the NLD'; and Moe Moe, 'Parliament approves funds for internet oversight body', *The Irrawaddy*, 21 March 2018, www.irrawaddy.com/news/parliament-approves-funds-internet-oversight-body.html.

fear that the government may use the team to monitor their activities, given their reliance on online spaces for organising and recruitment.[20]

This online monitoring demonstrates how defamation laws can set a clear political culture where silencing dissent is considered an expected and necessary element to the NLD's governance. The government's actions have created an atmosphere of caution among civil actors – its loudest critics. If they continue to have their desired effects, then it is unlikely the government will consider relaxing these restrictions.

Myanmar's cultural dynamics

Unless proponents for reform can mobilise wider support to amend Article 66(d), the relationship between the NLD and the public will pose a significant barrier to change. They already face difficulties because a strong distrust for media and activism has been brewing in Myanmar since the NLD took power.[21] These public sentiments are influenced by the rhetoric of Aung San Suu Kyi, who has shown little sympathy for those charged under Article 66(d) despite being a former political prisoner herself.[22] She has not given any press conferences since she began her position, nor has she met with any domestic media, thus hampering the publics ability to report effectively on local issues.[23] The relationship that the NLD has with the public means that its leaders can set cultural attitudes with their actions, knowing that the public is likely to adopt these sentiments due to its support for the Aung San Suu Kyi and her party.[24] The NLD's influence over the public's perception of the media therefore makes it difficult for civil actors to attract sympathy and mobilise support around the issues they champion like reforming Article 66(d).

While this distrust for the media and activists continues to permeate throughout Myanmar's society, civil actors still consider education to be a key element to their role. Proponents of reform believe that their role should not focus solely on legal reform, but on educating the public about the importance of freedom of expression and civil rights more broadly.[25] Because these concepts are still largely

---

[20] Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[21] Lisa Brooten, 'Burmese media in transition', *International Journal of Communication* 10 (2016): 190–92.

[22] Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[23] Ibid.

[24] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018; and Anonymous interview with an editor of English-language Myanmar publication. Interview by ANU students. In-person interview. Yangon, Myanmar, 18 July 2018.

[25] Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

unfamiliar to Myanmar's cultural fabrics, directors of Athan, a local non-government organisation that promotes freedom of expression, say it is difficult to identify cultural and linguistic translations for these terms in their work.[26] These barriers are familiar to many civil actors in Myanmar who must engage in vernacularisation to develop cultural understandings of such concepts.[27] For the directors of Athan, this forms a difficult but necessary aspect to their work.[28]

Mobilising support for greater freedoms of expression will not be possible without a cultural understanding of their importance and practice. Even so, Athan understand that the core barriers exist in Myanmar's political, not cultural, frameworks.[29] Because NLD's leaders have considerable influence over the population, they argue that their work must focus on political advocacy.[30]

Myanmar's policy frameworks

One area where civil actors believe they have the strongest possibility for reform is the duplication of the law regarding the chargeable offences within Article 66(d). The Penal Code 500 already outlines defamation in Myanmar, for example.[31] Proponents of reform argue that very few charges like those under the Article 66(d) have been used outside of the Telecommunications Law, despite the Penal Code being in place since British rule.[32] While the Penal Code does not refer to digital spaces explicitly, interpretations of the legal instrument would allow it to protect against defamation.[33] Proponents for reform argue that the Penal Code contains sufficient provisions for defamation and that Article 66(d) is being used as a political tool. Therefore, they argue, it should be reformed to reduce duplication and limit its use as a tool to stifle debate. While these arguments were used by civil actors to promote reform before the 2017 amendment, the superficial nature of that change indicates that the argument was not convincing enough to sway the NLD-led government.[34]

---

[26] Ibid.

[27] Lynette Chua, 'The vernacular mobilization of human rights in Myanmar's sexual orientation and gender identity movement', in *Law & Society Review* 49, no. 2 (2015): 308, 316.

[28] Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[29] Ibid.

[30] Ibid.

[31] 'Myanmar: Briefing paper on criminal defamation laws', International Commission of Jurists, accessed 29 July 2018, www.icj.org/myanmar-briefing-paper-on-criminal-defamation-laws/.

[32] Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[33] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018 and interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

[34] Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018 and interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

Even if Article 66(d) was reformed, there are sufficient legal frameworks that could be used to limit debate and expression in online spaces. A host of media professionals have already been arrested under laws including the Penal Code and the Official Secrets Act in Myanmar for allegedly publishing defamatory and incorrect statements about government officials.[35] The frequency of these restrictions within Myanmar's legal framework demonstrates that, while reforming Article 66(d) of the Telecommunications Law will be an important step to relaxing freedoms on media and expression, the issue is far more pervasive. Meaningful reform to relax Myanmar's restrictions on freedoms of media and expression would need to extend beyond Article 66(d) and would require a significant shift to the cultural and political attitudes within Myanmar.

Conclusion

This paper has drawn upon primary and secondary research to assess the possibility of meaningful reform to Article 66(d) of the Telecommunications Law. It has argued that there are three key barriers to change: Myanmar's political context, cultural dynamics and existing policy framework. While there are possible avenues for change – from the favouring opinions of NLD representatives expressed during parliament debates about the 2017 amendments, to the mobilisation of civil society and the duplicity of laws – it is unlikely that Article 66(d) will change in the current context. The paper has outlined a host of barriers, from the NLD-led government's claims of legitimacy and its relationship with the public, to regional trends for online restrictions and the public's distrust for the media, which pose significant barriers for meaningful reform.

The NLD's resistance to reforming Article 66(d) sets a dangerous norm where the silencing of dissent and criticism becomes essential to sustaining its governance. It indicates that they are unwilling to engage civil and media actors in Myanmar's democracy and discredit the crucial role that such actors have played in the country's liberalisation until this point. For a government still grappling with a political transition, the NLD should be wary of the impact this norm may have on sustaining its leadership.

Bibliography

Primary research

Anonymous discussion with political researcher during ANU course dinner. Naypyidaw, Myanmar, 10 July 2018.

Anonymous interview with an editor of English-language Myanmar publication. Interview by ANU students. In-person interview. Yangon, Myanmar, 18 July 2018.

[35] Lisa Brooten, 'Burmese media in transition', 185.

Anonymous interview with Myanmar media and political analyst. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

Anonymous interview with member of Myanmar Press Council. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

Anonymous interview with peace process consultant. Interview by ANU students. In-person interview. Yangon, Myanmar, 18 July 2018.

Interview with Aung Khant and Maung Saungkha. Interview by ANU students. In-person interview. Yangon, Myanmar, 17 July 2018.

Secondary research

Aung, Ei Cherry. 'Myanmar's parliament debates controversial telecom law'. *Mizzima*, 29 July 2017. www.mizzima.com/news-features/myanmar%25E2%2580%2599s-parliament-debates-controversial-telecom-law.

Brooten, Lisa. 'Burmese media in transition'. *International Journal of Communication* 10 (2016): 182–99.

Cain, Geoffrey. 'Kill one to warn one hundred'. *The International Journal of Press/Politics* 19, no. 1 (2013): 85–107. doi.org/10.1177/1940161213508814.

Chua, Lynette. 'The vernacular mobilization of human rights in Myanmar's sexual orientation and gender identity movement'. *Law & Society Review* 49, no. 2 (2015): 299–332. doi.org/10.1111/lasr.12135.

Doherty, Ben. 'Silence of the dissenters'. *The Guardian*, 22 October 2010. www.theguardian.com/technology/2010/oct/21/internet-web-censorship-asia.

Farrelly, Nicholas and Win, Chit. 'Inside Myanmar's turbulent transformation'. *Asia & the Pacific Policy Studies* 3, no. 1 (2016): 38–47. doi.org/10.1002/app5.124.

Free Expression Myanmar. 'Myanmar's media freedom at risk'. Report, Yangon, Myanmar, 2018.

Freedom House. 'Freedom on the net 2016: Thailand country profile'. Accessed 28 July 2018. freedomhouse.org/report/freedom-net/2016/Thailand.

Hutchinson, M., Schiano, S. and Whitten-Woodring, J. 'When the fourth estate becomes a fifth column'. *The International Journal of Press/Politics* 21, no. 2 (2016): 165–87. doi.org/10.1177/1940161216632362.

Hutt, David. 'Is Facebook helping Vietnam suppress online dissent?'. *Asia Times*, 2 June 2018. www.atimes.com/article/is-facebook-helping-vietnams-pro-democracy-crackdown/ (site discontinued).

International Commission of Jurists. 'Myanmar: Briefing paper on criminal defamation laws'. Accessed 29 July 2018. www.icj.org/myanmar-briefing-paper-on-criminal-defamation-laws/.

McCarthy, Gerard. 'Cyber-spaces'. In *Routledge Handbook of Contemporary Myanmar*, edited by Adam Simpson, Nicholas Farrelly and Ian Holliday, 92–105. New York: Routledge, 2018.

Moe, Moe. 'Parliament approves funds for internet oversight body'. *The Irrawaddy*, 21 March 2018. www.irrawaddy.com/news/parliament-approves-funds-internet-oversight-body.html.

Myoe, Maung Aung. 'The NLD and Myanmar's foreign policy'. *Journal of Current Southeast Asian Affairs* 36, no. 1 (2017): 89–121. doi.org/10.1177/186810341703600104.

SayNoto66d. 'Home page'. Last modified 8 March 2018. Accessed 28 July 2018. www.saynoto66d.info/.

Spencer, Oliver, Oo, Zar Chi and Thein, Yin Yadanar. 'The 66(d) amendment: tinkering at the edges'. *Frontier Myanmar*, 15 September 2017. frontiermyanmar.net/en/the-66d-amendment-tinkering-at-the-edges.

Tea Circle Oxford. 'Myanmar's freedom of expression as a broken promise of the NLD'. 30 May, 2018. teacircleoxford.com/2018/05/30/myanmars-freedom-of-expression-as-broken-promise-of-nld/.

Yangyue, Liu. 'Transgressiveness, civil society and internet control in Southeast Asia'. *The Pacific Review* 27, no. 3 (2014): 383–407. doi.org/10.1080/09512748.2014.909520.

Rotten spaces: The role of waste in constructing the imaginal of the occupation in East Jerusalem

SARAH BARRIE

Abstract

Waste is an evocative aesthetic phenomenon. It is capable of producing a myriad of subliminal messages that permeate through the public space it defiles. During its occupation of East Jerusalem, Israel's discriminatory waste management policy assumes a particularly powerful role in transforming the experience of the public space. Indeed, it distorts and 'dirties' the experience of the city and consolidates the imaginal of the occupation of East Jerusalem. It does so by evoking an array of ideas that align with one's response to, and rejection of, waste. Consequently, this paper is an interrogation of how this aesthetic experience produced by waste is achieved through interweaving two narratives that are concurrently performed in this public space: the image of Palestinians as a rejected matter within the occupied East Jerusalem, and of Israel's defiance of international legal convention. These interwoven narratives operate to conjure a powerful image of the burgeoning potency of Israel's sovereign power in occupied East Jerusalem.

Introduction

There is perhaps no image as evocative of the ideas of dirt, decay and rejection as that of waste. The presence of waste within a public domain can consequently create a myriad of subliminal messages that permeate, much like the fetid stench of rotting waste, through a communal space. This occurs when waste becomes an aesthetic phenomenon. Such a phenomenon currently exists in occupied East Jerusalem, a city where waste performs a perverse aesthetic function that transforms the daily experience of Palestinians.[1] Indeed, in this space the presence of waste contributes to the construction of the image of occupation by virtue of the socio-legal messages defilement and decay convey.[2] As a

[1] Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018) 13–14.

[2] Patricia Branco and Richard Mohr, 'Odore di Napoli: What if Jurisprudence Came to us through Smell?' (Online Working Paper, The Westminster Law and Theory Lab, April 2015) 59–61.

result, the legal status of East Jerusalem as an occupied space metamorphoses into an aesthetic phenomenon, performed by the prevalence of waste within the public space. This prevalence is an occurrence attributable to an inadequate waste management policy, which is intended to regulate how waste is treated and disposed of.[3]

This paper consequently interrogates how Israel's discriminatory waste management policy, which renders East Jerusalem 'dirtied', contributes to the construction of the imaginal of the occupation within this public space. In doing so, this paper is divided into two parts: first, a discussion of the imaginal of the public space of occupied East Jerusalem, followed by an analysis of how the presence of waste within this space performs an aesthetic function to distort the experience of East Jerusalemites and construct the image of the occupation. As this paper opines, this aesthetic experience is achieved through interlaying two narratives that are performed by the presence of waste within East Jerusalem. These interwoven narratives represent the disparate images of Palestinians as a rejected matter within this occupied space and of Israel's defiance of international legal convention, conjuring a powerful image of the potency of Israel's sovereign power in East Jerusalem.

Part One: The imaginal of East Jerusalem

A brief account of the current legal and political status of Jerusalem is critical to comprehending the politics of waste management in East Jerusalem and the aesthetic function it achieves. Jerusalem is a city central to any narrative of Palestine and the Palestinian people.[4] It holds important political, religious, historical and cultural significances for both the Palestinian and Israeli people.[5] Israel has maintained 'effective control' of East Jerusalem since the 1967 Six-Day War, applying its 'law, jurisdiction and administration' over this part of the city.[6] Prior to Israel's military victory in 1967, East Jerusalem was under Jordanian administration and fell beyond the 'Green Line' of the 1949 Armistice. This Armistice had carved up the Israeli–Palestinian territories following the earlier 1948 war, known to Palestinians as the *Nakba* (catastrophe).[7] The year of 1967 therefore commenced the effective annexation of East Jerusalem to the broader Israeli municipality of Jerusalem. However, this annexation was not formally achieved by Israel until July 1980,[8] the same year that Israel also elevated Jerusalem

---

[3] Israel Ministry of Environmental Protection, *Waste Treatment Methods* (2015) Israel Ministry of Environmental Protection <www.sviva.gov.il/English/env_topics/Solid_Waste/WasteTreatmentMethods/Pages/default.aspx>.

[4] Rashid Khalidi, *Palestinian Identity: The Construction of Modern National Consciousness* (Columbia University Press, 2010) 13.

[5] The Peace and Democracy Forum and Ir Amim, *Solid Waste Management Policy in the Jerusalem District* (September 2008) The Jerusalem Policy Forum 3 <www.pdf-palestine.org/wast.pdf>.

[6] Eyal Benvenisti, *The International Law of Occupation* (Oxford University Press, 2nd ed, 2012) 203–4.

[7] Ibid.

[8] Lex Takkenberg, *The Status of Palestinian Refugees in International Law* (Clarendon Press, 1998) 211.

to the status of Israel's capital under its 1980 Basic Law.[9] Israel has since exercised its sovereign rights over East Jerusalem.[10] This annexation 'dismembers the West Bank'[11] by virtue of its division of East Jerusalem from the remaining Palestinian territory. East Jerusalem was further physically severed from the West Bank by virtue of the Israeli-constructed separation barrier built in 2002.[12]

The imaginal and the public space

A public space is the theatre of the imaginal and therefore the geographic space of East Jerusalem has become the grand stage for the performance of Israel's sovereign power. This is because power is a 'fundamentally spatial phenomenon'[13] that is legitimised in the public space. The 'imaginal' is therefore the aesthetic dimension of space,[14] and is a reflection of how a sovereign intends its image of power will permeate through, and be experienced in, a public space by the body politic.

A sovereign is consequently 'nothing but an imaginal being';[15] it relies on its imaginal constitution to legitimise its power.[16] A sovereign does so by using a public space to 'disrup[t] the relationship between the visible, the sayable, and the thinkable without having to use the terms of a message as a vehicle'.[17] Indeed, a sovereign requires the body politic to visualise is subordination to its power, without being directly told this. As a result, how a sovereign arranges aesthetic space affects one's experience of power and legal meaning within the public domain.[18]

Occupation

Israel maintains effective control of the arrangement of aesthetic space within East Jerusalem by virtue of its occupying power and lack of conformity with the corresponding laws that apply to its occupation. Contrary to Israel's argument that the annexation of East Jerusalem was a purely administrative measure, the international community considers East Jerusalem to be occupied under international law.[19] As a result, in theory several international legal instruments apply to Israel's exercise of power

---

[9] Virginia Tilley (ed), *Beyond Occupation: Apartheid, Colonialism and International Law in the Occupied Palestinian Territories* (Pluto Press, 2012) 38.

[10] Ibid.

[11] Ibid.

[12] Michael Lynk, *Report of the Special Rapporteur on the Situation of Human Rights in the Palestinian Territories Occupied since 1967*, UN Doc A/71/554 (19 October 2016) 23.

[13] Gary Fields, *Enclosure: Palestinian Landscapes in a Historical Mirror* (University of California Press, 2017) 7.

[14] Chiara Bottici, *Imaginal Politics: Images Beyond Imagination and the Imaginal* (Columbia University Press, 2014) 96.

[15] Ibid.

[16] Ibid 93.

[17] Jacques Rancière, *The Politics of Aesthetics: The Distribution of the Sensible* (Continuum, 2004) 63.

[18] Ibid.

[19] Takkenberg, above n 8, 211.

in East Jerusalem. The most relevant legal instruments are the Fourth Geneva Convention of 1949[20] and the 1907 Hague Regulations.[21]

Despite Israel being a party to the Fourth Geneva Convention,[22] it contends that the Convention 'is not applicable *de jure* to the situation prevailing in the Occupied Palestinian Territory'.[23] Instead, Israel only accepts the '*de facto* application of what it calls the "humanitarian provisions" of the Convention'.[24] Conversely, while Israel is not formally bound by the 1907 Hague Regulations, Israel's Supreme Court has recognised the customary status of the Regulations and considers Israel domestically bound by its obligations.[25]

Irrespective of the contentious application of international occupation law to Israel's exercise of sovereign power, the United Nations Human Rights Council (UNHRC) argues that as an occupying force, Israel is bound by principles set out under international humanitarian law.[26] Under international law, the occupying power is considered the trustee of 'public order and civil life in the territory under its control' with the occupied population the 'beneficiaries of this trust'. [27] Consequently, subjugating an occupied population is a breach of this trust.[28] Specifically, article 47 of the Fourth Geneva Convention provides the protection that a territory and its population are not to be deprived of the rights and protections afforded by the Convention as a result of annexation.[29]

As Gross argues, there are two readings of the laws of occupation.[30] One reading is that it is a benevolent system that guarantees occupation 'will not be akin to conquest, colonialism, or apartheid but will rather be a temporary rule that will benefit the local population until the territory is freed.'[31] Conversely, a belligerent reading views occupation as a legally endorsed phenomenon, which profits from a 'cloak of

---

[20] *Geneva Convention Relative to the Protection of Civilian Persons in Time of War (Fourth Geneva Convention)*, opened for signature 12 August 1949, 75 UNTS 287 (entered into force 21 October 1950) ('Fourth Geneva Convention').

[21] *Hague Convention (IV) Respecting the Laws and Custom of War on Land and Its Annex: Regulations Concerning the Laws and Customs of War on Land*, opened for signature 18 October 1907, International Peace Conference 1907 (entered into force 26 January 1910) ('1907 Hague Regulations').

[22] Tilley, above n 9, 6.

[23] Peter Maurer, 'Challenges to International Humanitarian Law: Israel's Occupation Policy' (2012) 94(888) *International Review of the Red Cross* 1503, 1506.

[24] Ibid.

[25] Tilley, above n 9, 6.

[26] Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018) 2.

[27] Aeyal Gross, *The Writing on the Wall: Rethinking the International Law of Occupation* (Cambridge University Press, 2017) 18.

[28] Ibid.

[29] Ibid 24.

[30] Ibid 20.

[31] Ibid.

temporality and the stamp of international legality' because it is considered an 'accepted legal phenomenon' unlike apartheid or colonialism.[32] This paper opines that the occupation of East Jerusalem conforms to a belligerent reading of occupation law such that management of the territory is to the disadvantage of the local population. Indeed, the occupation has had a particularly detrimental effect on East Jerusalem since 1967,[33] and manifests as the 'recurring excesses of armed violence, the ensuring grief among the people affected, and the trauma among the broader community'.[34]

Space becomes another conquest of a sovereign's power that is transformed within the aesthetic realm to constitute legal authority. As a result, East Jerusalem has been progressively reshaped and reimagined under Israel's military occupation.[35] This public space has thus been distorted to become an imaginal manifestation of occupation. This is because space is not an 'absolute quantity' but is instead a 'human one; a function of the time that man uses to traverse it'. [36] Indeed, in East Jerusalem, one witnesses the phenomenon of space being reduced to a purely human construct; it does not have absolute limits, only those delineated by Israel to demarcate its power within the occupied space. As the next part of this paper argues, this conquest is achieved by the powerful aesthetic function of waste.

Part Two: The aesthetics of waste and the rotten public space of East Jerusalem

Waste management in East Jerusalem

Contrary to the protections afforded to an occupied population under international law, East Jerusalem is subject to a single discriminatory waste management policy regulated by the Jerusalem municipality.[37] The policy establishes that all waste from Jerusalem is to be collected and deposited in the Abu Dis landfill,[38] which is a neighbourhood beyond the separation barrier in the West Bank under Palestinian territorial authority.[39] This practice in and of itself violates international law which mandates that 'waste not be transmitted across borders without agreement'.[40]

---

[32] Ibid.

[33] Maurer, above n 23, 1507.

[34] Ibid 1505.

[35] Zahraa Zawawi and Eric Corjin and Bas Van Heur, 'Public Spaces in the Occupied Palestinian Territories' (2013) 78(4) *GeoJournal* 743, 743.

[36] Bottici, above n 14, 94.

[37] The Peace and Democracy Forum and Ir Amim, above n 5, 7 < www.pdf-palestine.org/wast.pdf>.

[38] Ibid.

[39] Ibid 6.

[40] Ibid 4.

The roots of discrimination against East Jerusalem Palestinians under this waste management policy are found in the allocation of the Jerusalem municipality's budget. According to the UNHRC, 'approximately 8 to 10 per cent of the budget of Jerusalem Municipality is allocated to Palestinians in East Jerusalem though they account for 37 per cent of the city's population.'[41] This practice occurs despite the fact Palestinians living in East Jerusalem pay the same taxes as all other residents.[42] Therefore, while this policy envisages that the Jerusalem municipality will collect all of East Jerusalem's waste, in practice this is irregular and inadequate.[43]

The practices of Palestinians in light of this discriminatory policy have significant ramifications for the environment and health of the citizens. Indeed, an unknown amount of waste is disposed of illegally or burned in East Jerusalem,[44] and the public space is littered with 'uncollected waste spread on the streets, overflowing dumpsters, and informal dumping'.[45] In addition to irregular garbage collection in East Jerusalem, the tipping fees are exorbitant, which creates 'an economic incentive for poorer communities to dump illegally'[46] and waste is frequently burned.[47] These practices increase the presence of toxins in the air and the contamination of water supplies,[48] 'severely affect[ing] the quality of life of the Palestinians residing in East Jerusalem.'[49] There is no enforcement of Israeli municipal regulations that cover waste management in East Jerusalem.[50]

The aesthetics of waste

Waste is a potent aesthetic phenomenon. The presence of waste and filth within a public space is 'never a unique isolated event'; it occurs where there is a hierarchy of ideas and a system of social order.[51] Indeed, waste is considered to be the 'matter out of place'[52] within a system. This is in part due to our

---

[41] Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018) 6.

[42] United Nations Relief and Works Agency, *Barrier Impacts on Waste Management* (2012) United Nations Relief and Works Agency 1 <www.unrwa.org/userfiles/20120628121143.pdf>.

[43] Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018) 6.

[44] The Peace and Democracy Forum and Ir Amim, above n 5, 9 <www.pdf-palestine.org/wast.pdf>.

[45] United Nations Relief and Works Agency, *Barrier Impacts on Waste Management* (2012) United Nations Relief and Works Agency 1 <www.unrwa.org/userfiles/20120628121143.pdf>.

[46] The Peace and Democracy Forum and Ir Amim, above n 5, 9 <www.pdf-palestine.org/wast.pdf>.

[47] Ibid 10.

[48] Ibid.

[49] Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018) 6.

[50] The Peace and Democracy Forum and Ir Amim, above n 5, 4 <www.pdf-palestine.org/wast.pdf>.

[51] Mary Douglas, *Purity and Danger: An Analysis of Concepts of Pollution and Taboo* (Routledge & Kegan Paul, 1966) 35–41.

[52] Ibid 35.

instinctive rejection of appearances of filth, disorder and waste, and linking these ideas to concepts of decreased functionality.[53]

The dirtying of East Jerusalem by waste is a regular occurrence in the daily lives of Palestinians. In East Jerusalem, poor waste management policies result in skips filled with rubbish littering the streets, overflowing with garbage dumped beside them. In the summer months, the scent of rotting food diffuses through the streets and in some neighbourhoods the 'smell of burned garbage hovers over the narrow streets.'[54] This is contrasted with Israeli West Jerusalem, and the settlements that have been built in the East, which are pristine as a result of regular waste collection and municipal workers frequently tidying the streets.

Consequently, waste is never *just* waste in East Jerusalem as its presence upholds ideas of the disorder and defilement of the public space. As Saito argues, in determining whether something appears messy, the preoccupation is not necessarily with the item but with the object's displacement.[55] Indeed, waste stored and disposed of correctly is not a particularly messy phenomenon; rubbish is not necessarily dirty unless it is defiling a public space, or dirtying what should otherwise be clean.[56] Thus, the purposive dirtying of East Jerusalem, as a consequence of Israel's discriminatory waste management policy, intends to render this public space messy, chaotic and defiled. Further, given that concepts of dirt and mess are 'context-dependent and culturally constructed',[57] the symbolic loading of dirt within East Jerusalem is amplified owing to the 'hygienic phobia of Zionism'.[58]

Similar to the images of dirt and waste, associated smells can organise experiences and connections to places. For example, in a study of experiences of the smells in the city of Naples, the concepts of criminality, harsh reality and despair were evoked by smells of blood and 'defiance'.[59] Applied to this context, it is possible to argue that ideas of subordination, abjection and rejection are evoked by the scents of rotting food and scorched waste. However, the experience of smell within a city differs between those who are considered cultural insiders and outsiders.[60]

Here it can be said that Palestinians 'inhabit the smells of their city, which in turn regulate their lives' whereas outsiders, such as Israelis or foreigners 'notice the transgressive smells of garbage, [and] the

---

[53] Yuriko Saito, *Everyday Aesthetics* (Oxford University Press, 2007) 159.

[54] Udi Shaham, 'Shuafat: The Refugee Camp in the Heart of Jerusalem', *The Jerusalem Post* (online) 23 December 2017 <www.jpost.com/Opinion/Shuafat-The-refugee-camp-in-the-heart-of-Jerusalem-519731>.

[55] Saito, above n 53, 156.

[56] Ibid 155.

[57] Ibid 154.

[58] Eyal Weizman, *Hollow Land: Israel's Architecture of Occupation* (Verso, 2007) 20.

[59] Branco and Mohr, above n 2, 69.

[60] Ibid 64.

pathological threat of the foreign unclean.'[61] Therefore, smells are capable of producing powerful aesthetic messages within East Jerusalem that distinguish cultural insiders and outsiders' responses to this space. As a result, from an outsider's perspective, the aesthetic effect of this phenomenon is that waste 'dirties' the imaginal of East Jerusalem and creates a rotten public space.

Rotten space: The aesthetic experience of East Jerusalem

The aesthetics of waste in East Jerusalem consequently produces two interwoven and colliding narratives of micro and macro imaginaries within this public space. These two narratives are derived from Douglas' account of dirt, such that '[o]ur idea of dirt is compounded of two things, care for hygiene and respect for convention'.[62] The first story of potentiality within this space is the rejection of the Palestinian population as 'matter out of place'. The other, occurring at the macro level, is Israel's disregard for international law and convention. As de Sousa Santos argues, 'socio-legal life is constituted by different legal spaces operating simultaneously on different scales and from different interpretative standpoints.'[63] Therefore, the convergence of these two 'symbolic universes'[64] produces an interwoven socio-legal narrative of occupation and Israel's burgeoning sovereign power within East Jerusalem.

Palestinians as 'matter out of place'

From an outside perspective, poor waste management in East Jerusalem embodies a disregard for hygiene and affirms Palestinian 'otherness' owing to Zionist dirt phobia. Taking Douglas' approach, the presence of waste in the public space of East Jerusalem is therefore symbolic of the social subordination of Palestinians.[65] In this regard, Weizman argues that the presence of waste and sewerage within this space 'affirms a common national territorial imagination that sees the presence of Palestinians as "defiled" substance within the "Israeli landscape" or as "matter out of place"'.[66]

Building on from this narrative of 'matter out of place', Israel's waste management policy becomes a coercive measure to encourage Palestinians out of East Jerusalem. This is achieved through the everyday experience of living in a defiled environment that appears subordinate to its Western counterpart. As Saito argues, pleasure derived from an aesthetic experience is quashed when the experience is chaotic or messy.[67] Therefore, in East Jerusalem waste operates to reduce the pleasure

---

[61] Ibid 70.

[62] Douglas, above n 51, 7.

[63] Boaventura de Sousa Santos, 'Law: A Map of Misreading. Toward a Postmodern Conception of Law' (1987) 14(3) *Journal of Law and Society* 279, 288.

[64] Ibid 286.

[65] Douglas, above n 51, 2.

[66] Weizman, above n 58, 20.

[67] Saito, above n 53,156.

derived from experiencing the space. As a result, the UNHRC has noted that practices such as poor waste management 'create a coercive environment that places Palestinians under pressure to leave'.[68] Further, in the context of poor sewerage in the West Bank, Weizman argues that it 'marks the point of collision between two meanings – a metaphorical political notion concerned with the health of the state, and the literal physical sensation of abjection.'[69]

A similar phenomenon occurs in East Jerusalem, whereby the presence of waste emphasises the poor health of the city and the abjection of everyday life for Palestinians. However, as Branco and Mohr explain, cultural insiders and outsiders have disparate reactions to garbage and filth within a public space. They argue that 'the outsiders find regularities, but not rules.'[70] The outside observer 'does not necessarily have to accept the rules of a legal system, while the internal point of view is that of the members of a group who are governed by the rules of the legal system and who accept these rules as standards of conduct.'[71] Therefore, this outsider perspective does not consider that presence of waste constitutes a legal order, but instead regularises the concept of Palestinians as matter out of place. Conversely, perhaps from an internal perspective, waste management is a phenomenon governed purely by law that does not defile or reduce the social position of Palestinians. This is because, as Branco and Mohr aver, '[p]eople living "inside" a particular legal regime tend not to think about the law, as an abstraction: it is part of everyday life.'[72]

Nevertheless, the aesthetics of a public space are pivotal to producing an image of the community that inhabits it. Indeed, as Bottici opines '[a] community is not a community until it is gathered together and unified in a pictorial (re)presentation.'[73] Thus here, at least from a culturally 'outside' perspective, the imaginal of the public space works to represent the Palestinians as a community out of place, and subjugated within this occupied territory. This is achieved by virtue of the fact that meanings found in landscapes and the narratives they produce of their inhabitants are 'anchored … on the basis of viewers' interpretations of what they are observing'.[74]

Israel's defiance of convention

The narrative that converges with the above story is that of Israel's defiance of convention and flagrant disregard for the illegality of its activities within occupied East Jerusalem. As Bottici avers, '[p]olitics

---

[68] Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018) 14.

[69] Weizman, above n 58, 20.

[70] Branco and Mohr, above n 2, 71.

[71] Ibid.

[72] Ibid.

[73] Bottici, above n 14, 91.

[74] Fields, above n 13, 7.

has to adorn itself with the dress of legitimacy; it needs its own apparatus of glory; otherwise it accounts to sheer violence.'[75] The sovereign does so within the imaginal, as politics requires a fabricated image of itself to legitimise its power. However, here we see the inversion of Bottici's aphorism.

The inversion of Bottici's aphorism occurs because the discriminatory nature of Israel's waste management policy translates into an aesthetic experience that emphasises the illegitimacy and illegal nature of Israeli rule in East Jerusalem. Indeed, the very application of Israel's legal system and policies to annexed East Jerusalem stands in contrast to international law.[76] Further, the sanitary issues presented by exposed, rotting waste violates the Fourth Geneva Convention, which sets out that the '[o]ccupying Power has the duty of ensuring and maintaining … hygiene in the occupied territory … [and take] preventive measures necessary to combat the spread of contagious diseases and epidemics'.[77]

The defilement of the public space thus affirms what is happening at a macro legal level: Israel is contravening principles of occupation law, including the protections that it is mandated to provide the Palestinian population.[78] This aligns with the belligerent reading of the laws of occupation, such that Israel is 'profiting from the cloak of temporality and the stamp of international legality'[79] in perpetuating a system that would otherwise be considered akin to conquest or apartheid. Therefore, waste management practices in East Jerusalem are used to both affirm Israel's sovereign power and to perform a legal narrative of Israel's illegitimate control, creating an everyday reminder of Israel's inviolable sovereign power. In this imaginal the sovereign is consequently a truly potent, petrifying and powerful beast.

Conclusion

Waste management has become crucial in creating the imaginal of the occupation in East Jerusalem through the construction of a coercive environment in which Palestinians are 'matter out of place'[80] within a space of Israel's burgeoning sovereign power. Therefore, while in a legal sense, '[o]ccupation has traditionally been regarded as a factual matter',[81] in the context of East Jerusalem it becomes a remarkable aesthetic phenomenon. This phenomenon operates to particularly transform the cultural outsider's experience of East Jerusalem by producing a narrative of Palestinians' 'otherness' within the rotten public space in which they reside. Interwoven with this phenomenon is a vivid image of the

---

[75] Bottici, above n 14, 101.

[76] Maurer, above n 23, 1508.

[77] *Fourth Geneva Convention,* opened for signature 12 August 1949, 75 UNTS 287 (entered into force 21 October 1950) art 56.

[78] Maurer, above n 23, 1508.

[79] Gross, above n 27, 20.

[80] Douglas, above n 51, 35.

[81] Yael Ronen, 'Illegal Occupation and Its Consequences' (2008) 41(1) *Israel Law Review* 201, 201.

potency of Israel's sovereign power, which sanctions the defilement of this public space. The convergence of the two narratives within East Jerusalem consequently fabricates the insidious tapestry of socio-legal life under Israel's military occupation in a manner that can only be achieved through such an evocative aesthetic phenomenon as waste.

Acknowledgements

The author would like to thank Professor Desmond Manderson for his encouragement to pursue this study and for his review of this work.

Bibliography

Articles/books/reports

Arai-Takahashi, Yutaka, 'Preoccupied with Occupation: Critical Examinations of the Historical Development of the Law of Occupation' (2012) 94(885) *International Review of the Red Cross* 51

Azoulay, Ariella, 'Potential History: Thinking through Violence' (2013) 39(3) *Critical Inquiry* 584

Benvenisti, Eyal, *The International Law of Occupation* (Oxford University Press, 2nd ed, 2012)

Bottici, Chiara, *Imaginal Politics: Images Beyond Imagination and the Imaginal* (Columbia University Press, 2014)

Coon, Anthony, *Town Planning under Military Occupation: An Examination of the Law and Practice of Town Planning in the Occupied West Bank* (Dartmouth Publishing Company Limited, 1992)

de Sousa Santos, Boaventura, 'Law: A Map of Misreading. Toward a Postmodern Conception of Law' (1987) 14(3) *Journal of Law and Society* 279

Douglas, Mary, *Purity and Danger: An Analysis of Concepts of Pollution and Taboo* (Routledge & Kegan Paul, 1966)

Fields, Gary, *Enclosure: Palestinian Landscapes in a Historical Mirror* (University of California Press, 2017)

Gross, Aeyal, *The Writing on the Wall: Rethinking the International Law of Occupation* (Cambridge University Press, 2017) doi.org/10.1017/9781316536308

Khalidi, Rashid, *Palestinian Identity: The Construction of Modern National Consciousness* (Columbia University Press, 2010)

Listik, Yonathan, 'Aesthetic Regime's Occupation of Representation' (2018) 59(139) *Kriterion: Revista de Filosofia* 309

Maurer, Peter, 'Challenges to International Humanitarian Law: Israel's Occupation Policy' (2012) 94(888) *International Review of the Red Cross* 1503

Mitchell, WJT, 'Image, Space, Revolution: The Arts of Occupation' (2012) 39(1) *Critical Inquiry* 8

Rancière, Jacques, *The Politics of Aesthetics: The Distribution of the Sensible* (Continuum, 2004)

Ronen, Yael, 'Illegal Occupation and Its Consequences' (2008) 41(1) *Israel Law Review* 201

Saito, Yuriko, *Everyday Aesthetics* (Oxford University Press, 2007)

Shusterman, Richard and Adele Tomlin (eds), *Aesthetic Experience* (Routledge, 2008)

Takkenberg, Lex, *The Status of Palestinian Refugees in International Law* (Clarendon Press, 1998)

Tilley, Virginia (ed), *Beyond Occupation: Apartheid, Colonialism and International Law in the Occupied Palestinian Territories* (Pluto Press, 2012)

Weizman, Eyal, *Hollow Land: Israel's Architecture of Occupation* (Verso, 2007)

Zawawi, Zahraa and Eric Corjin and Bas Van Heur, 'Public Spaces in the Occupied Palestinian Territories' (2013) 78(4) *GeoJournal* 743

International material

G*eneva Convention Relative to the Protection of Civilian Persons in Time of War (Fourth Geneva Convention)*, opened for signature 12 August 1949, 75 UNTS 287 (entered into force 21 October 1950)

Hague Convention (IV) Respecting the Laws and Custom of War on Land and Its Annex: Regulations Concerning the Laws and Customs of War on Land, opened for signature 18 October 1907, International Peace Conference 1907 (entered into force 26 January 1910)

Human Rights Council, *Report of the United Nations High Commissioner for Human Rights on Israeli Settlements in the Occupied Palestinian Territory, including East Jerusalem, and the Occupied Syrian Golan*, UN Doc A/HRC/37/43 (6 March 2018)

Lynk, Michael, *Report of the Special Rapporteur on the Situation of Human Rights in the Palestinian Territories Occupied since 1967*, UN Doc A/71/554 (19 October 2016)

Other

Adamteva, *Waste Management in Shuafat, East Jerusalem,* Adamteva <www.adamteva.org.il/?CategoryID=1244&ArticleID=2075>

Branco, Patricia and Richard Mohr, 'Odore di Napoli: What if Jurisprudence Came to us through Smell?' (Online Working Paper, The Westminster Law and Theory Lab, April 2015)

Israel Ministry of Environmental Protection, *Waste Treatment Methods* (2015) Israel Ministry of Environmental Protection <www.sviva.gov.il/English/env_topics/Solid_Waste/WasteTreatmentMethods/Pages/default.aspx>

Shaham, Udi, 'Shuafat: The Refugee Camp in the Heart of Jerusalem', *The Jerusalem Post* (online) 23 December 2017 <www.jpost.com/Opinion/Shuafat-The-refugee-camp-in-the-heart-of-Jerusalem-519731>

The Peace and Democracy Forum and Ir Amim, *Solid Waste Management Policy in the Jerusalem District* (September 2008) The Jerusalem Policy Forum <www.pdf-palestine.org/wast.pdf>

United Nations Relief and Works Agency, *Barrier Impacts on Waste Management* (2012) United Nations Relief and Works Agency <www.unrwa.org/userfiles/20120628121143.pdf>

Contesting the universality of human rights for women: An examination of violence against women in the Global South and the Global North

JESSICA ELLIOTT

Abstract

Despite the purported universality of human rights, the human rights of women are *not* universal. The human rights of women are systematically violated across both the Global North and Global South in the form of pervasive violence against women. Using postcolonial feminism and intersectionality and two case studies of honour killings in India and domestic violence against women in the United States, this paper makes two arguments. First, this paper argues that irrespective of whether you are 'wealthy, white and from the west' in the Global North or the converse in the Global South, the human rights of women are systematically violated. Second, this paper argues that human rights violations against women in the Global North are ignored by a focus on violence against 'other' women in the Global South and minority communities in the Global North. This paper advocates the adoption of an intersectional approach to recognise that human rights are not universal for all women who are vulnerable to violence because of their gender.[1]

Introduction

Despite a concerted focus on incorporating women's rights into international human rights law, human rights are anything but universal for women. While women may have human rights, these human rights are systemically violated, irrespective of whether the woman is 'wealthy, white and western' in the Global North or poor, not-white and from the Global South.[2] This paper focuses on violence against

[1] This article was originally prepared by the author as an essay in response to the statement 'Human rights are anything but universal. If you're wealthy, white and from the west you have them. If you're not, you don't. Critically evaluate this statement, with reference to two human rights case studies'.

[2] Charlotte Bunch, 'Women's Rights as Human Rights: Toward a Re-vision of Human Rights;'(1990) 12(4) *Human Rights Quarterly* 486.

women (VAW), a universal phenomenon that violates the human rights of women across all cultures.[3] The United Nations Secretary-General has described VAW as 'one of the most heinous, systematic and prevalent human rights abuses in the world'.[4] I make two arguments by adopting a postcolonial feminist and intersectionality lens. First, I argue that, irrespective of whether you are 'wealthy, white and from the west' in the Global North or the converse in the Global South, the human rights of women are systemically violated. While women may *have* human rights, these human rights are far from guaranteed in reality. Second, I argue that not only are human rights not universal for women, the human rights violations of wealthy, white and Western women are ignored by a focus on violence against 'other' women in the Global South and minority communities in the Global North. A focus on VAW against women in the Global South and minority communities in the Global North fails to appreciate that VAW reflects patriarchal structural dynamics of power that subordinate women universally.

In my making my arguments, I analyse two manifestations of VAW: honour killings of women in India; and domestic violence perpetrated against women in the United States. These case studies enable a comparison of the international human rights treatment of VAW in the private sphere in the Global North and Global South. In both case studies, the human rights of women, whether 'brown, poor and from the Global South' or 'white, wealthy, and western', are violated. I draw attention to the human rights condemnation of honour killings as affecting 'other' women, while simultaneously rendering invisible domestic violence perpetrated against women in the US. I adopt an intersectional approach to examine the compounding nature of an individual's axes of differences in determining their experiences of human rights abuses.[5]

What is VAW?

Both domestic violence and honour killings are forms of VAW. I use the definition of VAW used in the Declaration on the Elimination of Violence against Women (DEVAW): 'any act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women'.[6] The definition of domestic violence is contested.[7] I define domestic violence broadly as 'the combination of physical and/or sexual violence with a variety of control tactics such as economic,

---

[3] Charlotte Watts and Cathy Zimmerman, 'Violence Against Women: Global Scope and Magnitude' (2002) 359(9313) *Lancet* 1232.

[4] United Nations Secretary-General, *Message on the International Day for the Elimination of Violence Against Women* (25 November 2007) <www.un.org/events/women/violence/2007/sg.shtml> (site discontinued).

[5] Kimberle Crenshaw, 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics' (1989) 129 *University of Chicago Legal Forum* 139.

[6] *Declaration on the Elimination of Violence against Women*, GA Res 48/104, UN GAOR, 48th sess, 85th plen mtg, UN Doc A/RES/48/104 (20 December 1993) art 1.

[7] Nicky Ali Jackson (ed), *Encyclopedia of Domestic Violence* (Taylor & Francis, 2007) 255.

emotional, social and spiritual abuse, the use of children and pets, and threats and intimidation'.[8] This definition reflects the complexity and multiplicity of forms of domestic violence.[9] I define honour killings as 'extreme acts of domestic violence culminating in the murder of a woman by her family or community'.[10] I appreciate that victims of violence may prefer the term 'survivor'. I use the term 'victim' to highlight the contrasting characterisation of some women as 'victims' in the Global South, while ignoring women who are subject to violence in the Global North. I also acknowledge that men are victims of domestic violence and honour killings.[11] However, I will focus exclusively on women due to limitations in the scope of this paper.

VAW is a violation of international human rights law

VAW has only recently been recognised as a violation of international human rights law. Despite the purported universality of international human rights law,[12] women have been historically neglected in international human rights law. Edwards notes that despite 'equality before the law and equal protection of the law being recognised as a right in all the major human rights treaties since 1945', VAW only gained prominence in the international community in the 1990s.[13] International human rights law has historically reflected the dichotomy between the public and private spheres by focusing on harms perpetrated in the public sphere.[14] Conversely, state interference is viewed as inappropriate in the private sphere, despite this being where the majority of human rights abuses against women, including VAW occur.[15]

The invisibility of women's issues in international human rights law prompted a movement to recognise 'women's rights as human rights', evidenced by the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW).[16] This movement led to significant advancements in the

---

[8] Sarah Wendt and Lana Zannettino, *Domestic Violence in Diverse Contexts* (Routledge, 2015) 2.

[9] Briana Barocas, Danielle Emery, Linda G. Mills, 'Changing the Domestic Violence Narrative: Aligning Definitions and Standards' (2016) 31(8) *Journal of Family Violence* 941, 941.

[10] Veena Meetoo and Heidi Safia Mirza, '"There is Nothing 'Honourable' about Honour Killings": Gender, Violence and the Limits of Multiculturalism' (2007) 30 *Women's Studies International Forum* 187, 187.

[11] See Babette C Drijber, Udo JL Reijnders and Manon Ceelen, 'Male Victims of Domestic Violence' (2013) 28(2) *Journal of Family Violence* 173; Phyllis Chesler, 'Worldwide Trends in Honor Killings' (2010) 17(2) *Middle East Quarterly Spring* 3, 5.

[12] Jack Donnelly, 'The Relative Universality of Human Rights' (2007) 29(2) *Human Rights Quarterly* 281.

[13] Alice Edwards, *Violence Against Women under International Human Rights Law* (Cambridge University Press, 2010) 7.

[14] Hilary Charlesworth, 'Are Women's Rights Human Rights?: International Law of Human Rights' (Speech delivered at the Alicia Johnson memorial lecture, Darwin, 3 September 1993).

[15] Hilary Charlesworth, 'Human Rights as Men's Rights' in Julie Peters and Andrea Wolper (eds), *Women's Rights, Human Rights: International Feminist Perspectives* (Routledge, 1995) 103, 103.

[16] Bunch, above n 1.

inclusion of women's rights in international human rights. CEDAW is the only Convention specifically addressed to women,[17] and has been described as 'universal in reach, comprehensive in scope and legally binding in character'.[18]

Notably, despite its alleged comprehensiveness, CEDAW notably does not explicitly prohibit VAW.[19] Charlesworth and Chinkin have criticised the failure of human rights instruments to specify VAW as a violation of human rights, as this failure reflects an implicit 'gendered hierarchy of rights' that continues to neglect the needs of women.[20] This has been remedied to some extent by DEVAW and CEDAW's General Recommendation 19.[21] General Recommendation 19 declares VAW as a form of sex discrimination under Article 1 of CEDAW.[22] General Recommendation 19 also specifies eight additional rights and freedoms impaired by VAW including, 'the right to life; the right not to be subjected to torture or to cruel, inhuman or degrading treatment or punishment [and] the right to liberty and security of person'.[23]

Challenges to the universality of human rights

The violation, not protection, of the human rights of women is universal. Despite the dramatic increase in the recognition of women and VAW in international human rights law, VAW remains a widespread human rights violation. The widespread and global nature of VAW challenges the purported universality of human rights. One in three girls and women will experience gender-related violence because they are a woman throughout their lifetime.[24] VAW is universal in nature because it is the product of the universal subordination of women.[25] Bunch attributes VAW to 'the structural relationships of power, domination and privilege between men and women in society. VAW is central

---

[17] *Convention on the Elimination of All Forms of Discrimination against Women*, opened for signature 18 December 1979, 1249 UNTS 13 (entered into force 3 September 1981) ('CEDAW'); Marsha A Freeman, Christine Chinkin and Beate Rudolf, 'Introduction' in Marsha A Freeman, Christine Chinkin and Beate Rudolf (eds), *The UN Convention on the Elimination of All Forms of Discrimination Against Women: A Commentary* (Oxford University Press, 2012) 453.

[18] Rebecca J Cook, 'Reservations to the Convention on the Elimination of All Forms of Discrimination against Women' (1990) 30 *Virginia Journal of International Law* 643, 643.

[19] Bonita Meyersfeld, *Domestic Violence and International Law* (Bloomsbury Publishing, 2010), 3.

[20] Hilary Charlesworth and Christine Chinkin, *The Boundaries of International Law: A Feminist Analysis* (Manchester University Press, 2000) 218.

[21] UN Committee on the Elimination of Discrimination Against Women, *CEDAW General Recommendations Nos. 19 and 20*, 11th sess, UN Doc A/47/38 (24 June 1992).

[22] Ibid [6].

[23] Ibid [7].

[24] United Nations Population Fund, *Gender-Based Violence* (2018) <www.unfpa.org/gender-based-violence>.

[25] Charlesworth, 'Are Women's Rights Human Rights?', above n 13.

to maintaining those political relations at home, at work and in all public spheres'.[26] As a result of the origins of VAW in patriarchal power structures, it is unsurprising that VAW is a human rights abuse that occurs in all countries in the world.[27]

Despite VAW occurring throughout the world, it is paradoxically characterised as a human rights abuse perpetrated against 'other' women in the Global South and minority communities in the Global North. The dichotomy between the Global North and Global South is used to differentiate the liberated 'white, wealthy and Western woman' in the Global North from the oppressed, agentless and vulnerable women in the Global South.[28] Mutua pithily states the 'face of the prototypical victim is non-white' and located in the Global South.[29] Conversely, the Global North is homogenously constructed as 'being friendly to human rights'.[30] Accordingly, advocates in the Global North portray themselves as 'rescuer for those in the South who are victims of human rights violations'.[31] This 'othering' of victims of VAW is exemplified in Spivak's renowned conception of postcolonial human rights interventions as 'white men saving brown women from brown men'.[32] This trope reflects the attitude that human rights abuses are perpetrated against 'other' women in the Global South, while ignoring the reality that all women may need saving from violence, including 'white women' from 'white men'.

Honour killings in India: A human rights violation of 'other' women

The international community and human rights organisations widely condemn honour killings in India.[33] Grewal, a scholar on media depictions of honour killings in India and the US states that 'honour killings' have a 'strong connotation – whether it is in the media, by NGOs or by governments; an "honour killing" is something that is understood to be a terrible injustice, a human rights violation and

---

[26] Charlotte Bunch, *Passionate Politics Essays 1968–1986: Feminist Theory in Action* (St Martin's Press, 1987) 491.

[27] Jennifer L Ulrich, 'Confronting Gender-Based Violence with International Instruments: Is a Solution to the Pandemic Within Reach?' (2000) 7(2) *Indiana Journal of Global Legal Studies* 629, 631.

[28] Karen Morgaine, 'Domestic Violence and Human Rights: Local Challenges to a Universal Framework' (2007) 34 *Journal of Sociology & Social Welfare* 109, 110; Raka Ray, 'Postcoloniality and the Sociology of Gender' in James W Messerschmidt et al (eds), *Gender Reckonings: New Social Theory and Research* (New York University Press, 2018) 73, 80.

[29] Makau Mutua, 'Savages, Victims, and Saviours: The Metaphor of Human Rights' (2001) 42(1) *Harvard International Law Journal* 201, 230.

[30] Inderpal Grewal, '"Women's Rights as Human Rights': Feminist Practices, Global Feminism, and Human Rights Regimes in Transnationality' (1999) 3(3) *Citizenship Studies* 337, 340.

[31] Morgaine, above n 27, 123.

[32] Gayatri Chakravorty Spivak, 'Can the Subaltern Speak?' in C Nelson and L Grossberg (eds), *Marxism and the Interpretation of Culture* (University of Illinois Press, 1988) 271, 297.

[33] Inderpal Grewal, 'Outsourcing Patriarchy' (2013) 15(1) *International Feminist Journal of Politics* 1, 2.

gender subordination'.[34] It is difficult to accurately document honour killings due to chronic under-reporting by families and weaknesses in data collection.[35] However, the India Democratic Women's Association estimates that 1,400 honour killings occur annually.[36]

Honour killings are perpetrated despite India's ratification of CEDAW and implementation of domestic legislation criminalising VAW.[37] The concluding observations of the combined fourth and fifth periodic reports of India by the Committee on the Elimination of Discrimination expressed concern about the 'persistence of so-called "honour crimes" perpetrated by family members against women and girls'.[38] This indicates that although Indian women may have both international and domestic rights to be free from violence, this does not guarantee the protection of their rights against violence in reality.

Condemnation of honour killings emphasises the cultural roots of this form of VAW.[39] Metoo and Mirza note that reports of honour killings in Western media are 'often sensationalist, and engage in cultural stereotyping which puts the gaze on the "other"'.[40] This is apparent in the three resolutions adopted by the United Nations General Assembly that exclusively target 'crimes against women committed in the name of honour', rather than opposing VAW more generally.[41] The term 'honour' conjures preconceived notions of cultural difference of the Global North to 'other' societies with traditional honour codes.[42] Furthermore, the term 'honour' unnecessarily distinguishes this 'killing' from other types of VAW that also occasion death.[43] This is not to suggest that there are not cultural elements to the specific manifestation of forms of VAW.[44] I recognise that there is a diversity of women's experiences of violence both within and across cultures. I suggest, however, that it is a mistake to focus exclusively on cultural motivations of honour killings.

---

[34] Ibid.

[35] Holly Johnson, Natalia Ollus and Sami Nevala, *Violence Against Women: An International Perspective* (Springer, 2007) 10.

[36] Rao Arif Ali Khan, *Honour Killing: Roots and Remedies* (Mittal Publications, 2012) 114.

[37] Ibid.

[38] UN Committee on the Elimination of Discrimination against Women, *Concluding Observations on the Combined Fourth and Fifth Periodic Reports of India*, UN Doc CEDAW/C/IND/CO/4-5 (24 July 2014) [10].

[39] Rupa Reddy, 'Domestic Violence or Cultural Tradition? Approaches to "Honour Killing" as Species and Subspecies in English Legal Practice' in Aisha K Gill, Carolyn Strange and Karl Roberts (eds), *'Honour' Killing and Violence: Theory, Policy and Practice* (Palgrave Macmillan, 2014) 27.

[40] Meetoo and Mirza, above n 9, 194.

[41] Aisha K Gill, 'Introduction: "Honour" and "Honour"-Based Violence: Challenging Common Assumptions' in Aisha K Gill, Carolyn Strange and Karl Roberts (eds), *'Honour' Killing and Violence: Theory, Policy and Practice* (Palgrave Macmillan, 2014) 1, 12.

[42] Grewal, above n 29.

[43] Ibid.

[44] Gill, above n 40, 9.

The first reason is that a focus on cultural explanations of honour killings ignores how honour killings are a manifestation of the universal phenomenon of VAW. The characterisation of honour killings in terms of the 'other' ignores that honour killings, like all forms of VAW, reflect patriarchal systems of power that subordinate women. This view problematises Indian culture as the cause of honour killings, rather than recognising the origins of honour killings in the subordination of women. Furthermore, this view renders invisible the reality that VAW that occasions death is not restricted to cultures with honour codes in the Global South.

The failure to recognise the origins of honour killings in the structural subordination of women is apparent in the Special Rapporteur on Violence against Women's statement that 'manifestations of violence against women are a reflection of the structural and institutional inequality that is a reality for most women in India'.[45] Notably, the Special Rapporteur does not rely on cultural explanations for VAW in India. However, the Special Rapporteur fails to appreciate the universality of the 'structural and institutional inequality that is reality for most women' in the world, rather than exclusively women in India.

The second reason that cultural explanations of honour killings is inadequate is that it essentialises Indian women by portraying them as inherently vulnerable. Mohanty, a postcolonial feminist, has criticised Western feminist engagement with women in the Global South as producing a 'composite, singular "third world woman"',[46] characterised by '"third world difference" – that stable, ahistorical something that apparently oppresses most if not all the women in these countries'.[47] This construction of the oppressed 'third world woman' is exemplified in the characterisation of honour killings as perpetrated against the monolithic and agentless 'other' women. This essentialist view denies the agency of all Indian women by casting them as vulnerable, oppressed and in need of saving from their dangerous culture. This characterisation of Indian women as in need of saving is starkly contrasted with the portrayal of female victims of VAW in the US.

Domestic violence in the US: A human rights violation of 'white, wealthy and western' women

Despite the US' portrayal of itself as a leader in human rights, human rights abuses are routinely, yet invisibly, perpetrated against its female citizens. Like the female victims of honour killings in India, the human rights of 'white, wealthy and western' women in the US are also violated by domestic violence.

---

[45] Rashida Manjoo, *Report of the Special Rapporteur on Violence against Women, its causes and consequences*, UN Doc A/HRC/26/38/Add.1 (1 April 2014) [7].

[46] Chandra Talpade Mohanty, 'Under Western Eyes: Feminist Scholarship and Colonial Discourses' in Chandra Talpade Mohanty, Ann Russo and Loursdes Torres (eds), *Third World Women and the Politics of Feminism* (Indiana University Press, 1991) 51, 53.

[47] Ibid, 54.

As with honour killings, the prevalence of domestic violence is difficult to document. Nearly 20 people per minute experience domestic violence in the United States, equating to over 10 million victims annually.[48] Four out of five victims of domestic violence are women.[49] It is estimated that 23 per cent of women in the US are the victim of severe physical violence by an intimate partner.[50] I note that these statistics do not discern what percentage of victims were 'white, wealthy and western', and that domestic violence is more prevalent in immigrant communities in the US.[51] However, this paper contends that domestic violence in the US, like VAW more generally, does not discriminate between women. Being 'white, wealthy and western' does not prevent a woman from being a victim of domestic violence.

The limited visibility of domestic violence in the US reflects the characterisation of the US as a champion of human rights. Critics may argue that domestic violence in the US is not subject to the same hypervisibility as honour killings in India because domestic violence is a less 'egregious' human rights abuse than honour killings. However, this view fails to recognise that women die from domestic violence. The United Nations Secretary-General's first comprehensive report on VAW identified that 40 to 70 per cent of female murder victims in the US were killed by their husbands or male partners.[52] Further, the history, identity and values of the US are grounded in its characterisation as a 'site of freedom and human rights'.[53] However, this unequivocal characterisation of the US as a protector of human rights both ignores and renders invisible ongoing human rights violations perpetrated against women in the US. This ensures that human rights advocacy in the US typically focuses on human rights abuses in other countries rather than domestically. Soohoo, Albisa and Davis aptly note that, 'claims of human rights violations were levelled *by*, not *at*, the U.S. government'.[54]

The continuing association of VAW with 'other' women perpetuates the tendency to attribute cultural reasons as the cause of domestic violence, rather than questioning the structures of power which subordinate women globally. When domestic violence is acknowledged in the US, it is typically viewed as perpetrated exclusively or near-exclusively against women in minority non-white communities.[55] I

---

[48] National Coalition against Domestic Violence, *Statistics* (2015) <ncadv.org/statistics>.

[49] National Domestic Violence Hotline, *Get the Facts & Figures* (2018) <www.thehotline.org/resources/statistics/>.

[50] Ibid.

[51] Gretchen E Ely, 'Domestic Violence and Immigrant Communities in the United States: A Review of Women's Unique Needs and Recommendations for Social Work Practice and Research' (2004) 7(4) *Stress, Trauma, and Crisis* 223.

[52] United Nations Secretary-General, *In-Depth Study on All Forms of Violence against Women*, UN Doc A/61/122/Add.1 (6 July 2006) [115].

[53] Grewal, above n 29, 340.

[54] Cynthia Soohoo, Catherine Albisa and Martha F Davis (eds), *Bringing Human Rights Home: A History of Human Rights in the United States* (University of Pennsylvania Press, 2008) vii (emphasis in original).

[55] Leti Volpp, 'On Culture, Difference and Domestic Violence' (2002) 2 (11) *Journal of Gender, Social Policy & the Law* 393, 394.

appreciate that women in minority communities experience higher rates of domestic violence.[56] The association of domestic violence with non-white communities in the US is a subtle manifestation of the emphasis on cultural justifications for VAW perpetrated against 'other' women.[57] This perpetuates the narrative that domestic violence in the US is aberrant from, rather than reflective of, Western values. This juxtaposes the characterisation of honour killings as reflecting a dangerous Indian culture.

The US' human rights engagement therefore mirrors Spivak's colonial paradigm of 'white men saving brown women from brown men'. The US recognises that human rights abuses are perpetrated against 'brown women' by 'brown men', either in the Global South or minority communities in the Global North. This ignores that the human rights of women across all cultures are violated by VAW. The notion of 'white men saving *white* women from *white* men' would undermine the depiction of the US as a champion of human rights by highlighting human rights abuses against 'white, wealthy, western' women.

The failure of the US to ratify CEDAW ironically exemplifies the strength of its perception of itself as a champion of human rights. Women in the US are perceived as not 'needing' international human rights protection under CEDAW. The US is one of only six United Nations member states that has not ratified or acceded to CEDAW.[58] This attitude is exemplified by Mandhane's statement that although the 'United States has not yet ratified CEDAW … women in the U.S. enjoy substantial rights due to the non-discrimination provisions in the United States Constitution'.[59] This view fails to appreciate that these constitutional rights do not necessarily translate into the protection of those 'substantial rights' and ratification of CEDAW would further enshrine a culture of domestic human rights protections.

A focus on the professional successes of Western women ignores the reality that women are frequently subject to violence and sexual harassment in the workplace. Cohen notes that, '[i]n contrast with the dire condition of women in developing countries, the condition of women in Western countries is now outstanding'.[60] To support her argument, Cohen cites the rates of women at university and in managerial positions.[61] Cohen's characterisation of the 'condition of women in Western countries' as 'outstanding'

---

[56] Mutua, above n 28, 231.

[57] Deborah M Weissman, 'The Human Rights Dilemma: Rethinking the Humanitarian Project' (2004) 35 *Columbia Human Rights Law Review* 259, 316.

[58] Office of the High Commissioner for Human Rights, *Ratification status for CEDAW* <tbinternet.ohchr.org/_layouts/TreatyBodyExternal/Treaty.aspx?Treaty=CEDAW&Lang=en>.

[59] Renu Mandhane, 'The Use of Human Rights Discourse to Secure Women's Interests: Critical Analysis of the Implications' (2004) 10 *Michigan Journal of Gender & Law* 275, 310.

[60] Michelle Fram Cohen, 'The Condition of Women in Developing and Developed Countries' (2006) 11(2) *Independent Review* 261, 266.

[61] Ibid.

is problematic, however, because it ignores the epidemic nature of domestic violence and other forms of VAW. This issue is exemplified in the case of *Jessica Gonzales v United States*.

The landmark case of *Jessica Gonzales v United States* challenged the perception that the human rights of women in the US were always protected.[62] Unlike the US 'shining the spotlight' on other countries' human rights abuses, the case focused attention on the US' own human rights standards.[63] Gonzales was a victim of domestic violence whose children were killed after police failed to enforce a restraining order against her estranged husband. The case was the first instance in which a domestic violence victim brought a human rights action against the US. Many found the characterisation of the US as a human rights abuser disturbing.[64] A congressman told Gonzales, 'do you know how embarrassing it would be for an international body to call the US a violator of the rights of women and children?'[65] This statement illustrates the ongoing reluctance for the US to scrutinise the human rights of women in the US.

Significantly, the Inter-American Commission on Human Rights determined that the US had violated the rights of Gonzales.[66] The Inter-American Commission on Human Rights found that the US had failed to act with due diligence to protect Gonzales and her children from domestic violence, which violated the state's obligation 'not to discriminate and to provide for equal protection before the law', and 'their right to life under Article I of the American Declaration'.[67] It is therefore clear that the ongoing perception of the US as a leader of human rights ignores the reality that domestic violence remains a pervasive human rights abuse of women in the US.

An intersectional approach to victims of VAW

I advocate that the international human rights community should rely on Crenshaw's seminal theory of intersectionality to recognise that all women are vulnerable to violence because of their gender.[68] My case studies have challenged the current focus on human rights abuses perpetrated against 'other' women. Intersectionality challenges the crudeness of the classification that if 'you're wealthy, white and from the west you have [human rights]', and if 'you're not, you don't'. Instead, intersectionality recognises that individuals experience vulnerability on the basis of multiple compounding spheres of

---

[62] *Lenahan (Gonzales) v United States of America* (Inter-American Commission on Human Rights, Case No 12.626, 2011).

[63] Caroline Bettinger-Lopez, 'Jessica Gonzalez v United States: An Emerging Model for Domestic Violence and Human Rights Advocacy in the United States' (2008) 21 *Harvard Human Rights Journal* 183, 193.

[64] Ibid.

[65] Ibid.

[66] *Lenahan (Gonzales) v United States of America* (Inter-American Commission on Human Rights, Case No 12.626, 2011).

[67] Ibid, 5.

[68] Crenshaw, above n 4.

disadvantage, for example on the basis of gender, race, class and sexuality.[69] Intersectionality recognises that *all* women are vulnerable to violence because of their gender, as demonstrated in my two case studies. The human rights community can no longer continue to simply associate VAW with 'other' women while assuming that 'wealthy, white and western' women are invulnerable to human rights abuses.

Intersectionality also recognises that some women experience greater vulnerability to violence. For example, a non-white, poor woman in the Global South may be more vulnerable to violence because of the intersection of her gender in addition to the disadvantage afforded by her race, class and geographical location. Conversely, 'white, wealthy, and western' women may not experience spheres of disadvantage based on their race, class or geographical location. However, this does not mean that women who are 'white, wealthy and western' are never vulnerable to violence. All women are inherently vulnerable to violence because of their gender, in addition to potential other factors. For example, both the 'white, wealthy and western' woman and the 'non-white, poor and Global South' woman are more vulnerable to violence if they are LGBTIQ, rurally located, or disabled.[70] Therefore, an intersectional approach would foster a more nuanced understanding of the victims of domestic violence and disrupt the conceptions that only 'other' women in the Global South are vulnerable to VAW and that women in the West do not experience human rights violations.

Conclusion

This paper has challenged the universality of human rights for women by arguing that the human rights of women are systematically violated across both the Global North and Global South in the form of pervasive VAW. While the human rights of women may now be protected under international human rights law, this has not translated into the protection of human rights for women. My case studies of honour killings in India and domestic violence in the US have illustrated that the human rights of women are systemically violated. This occurs irrespective of whether the woman is 'wealthy, white and from the west' or 'poor, non-white and from the Global South'. Furthermore, I have established that the continuing focus on violence against 'other' women ignores the universal nature of VAW as an embodiment of patriarchal power structures that subordinate women. This denies the agency of women in the Global South while rendering human rights abuses perpetrated against 'white, wealthy and western' women invisible. It is imperative that an intersectional approach is adopted to recognise that human rights are not guaranteed for all women due to the interrelationship between the global patriarchy

[69] Ibid.

[70] Wendt and Zannettino, above n 7, 15.

and VAW. An intersectional approach would recognise the vulnerability of all women across both the Global North and Global South and facilitate the increased protection of the human rights of all women.

Acknowledgements

I thank Dr Jeremy Farrall, Associate Dean (Research) and Associate Professor of the ANU College of Law for his ongoing support.

Bibliography

Articles/books/reports

Barocas, Briana, Danielle Emery, Linda G Mills, 'Changing the Domestic Violence Narrative: Aligning Definitions and Standards' (2016) 31(8) *Journal of Family Violence* 941

Bettinger-Lopez, Caroline, 'Jessica Gonzalez v United States: An Emerging Model for Domestic Violence and Human Rights Advocacy in the United States' (2008) 21 *Harvard Human Rights Journal* 183

Bunch, Charlotte, 'Women's Rights as Human Rights: Toward a Re-vision of Human Rights' (1990) 12(4) *Human Rights Quarterly* 486

Bunch, Charlotte, *Passionate Politics Essays 1968–1986: Feminist Theory in Action* (St Martin's Press, 1987)

Charlesworth, Hilary, 'Human Rights as Men's Rights' in Julie Peters and Andrea Wolper (eds), *Women's Rights, Human Rights: International Feminist Perspectives* (Routledge, 1995) 103

Charlesworth, Hilary and Christine Chinkin, *The Boundaries of International Law: A Feminist Analysis* (Manchester University Press, 2000)

Chesler, Phyllis, 'Worldwide Trends in Honor Killings' (2010) 17(2) *Middle East Quarterly Spring* 3

Cohen, Michelle Fram, 'The Condition of Women in Developing and Developed Countries' (2006) 11(2) *Independent Review* 261

Cook, Rebecca J, 'Reservations to the Convention on the Elimination of All Forms of Discrimination against Women' (1990) 30 *Virginia Journal of International Law* 643

Crenshaw, Kimberle, 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics' (1989) 129 *University of Chicago Legal Forum* 139

Donnelly, Jack, 'The Relative Universality of Human Rights' (2007) 29(2) *Human Rights Quarterly* 281

Drijber, Babette C., Udo JL Reijnders and Manon Ceelen, 'Male Victims of Domestic Violence' (2013) 28(2) *Journal of Family Violence* 173

Edwards, Alice, *Violence against Women under International Human Rights Law* (Cambridge University Press, 2010) doi.org/10.1017/CBO9780511779671

Ely, Gretchen E, 'Domestic Violence and Immigrant Communities in the United States: A Review of Women's Unique Needs and Recommendations for Social Work Practice and Research' (2004) 7(4) *Stress, Trauma, and Crisis* 223

Freeman, Marsha A, Christine Chinkin and Beate Rudolf, 'Introduction' in Marsha A Freeman, Christine Chinkin and Beate Rudolf (eds), *The UN Convention on the Elimination of All Forms of Discrimination Against Women: A Commentary* (Oxford University Press, 2012)

Gill, Aisha K, 'Introduction: "Honour" and "Honour"-Based Violence: Challenging Common Assumptions' in Aisha K Gill, Carolyn Strange and Karl Roberts (eds), *'Honour' Killing and Violence: Theory, Policy and Practice* (Palgrave Macmillan, 2014) 1

Grewal, Inderpal, '"Women's rights as human rights": Feminist Practices, Global Feminism, and Human Rights Regimes in Transnationality' (1999) 3(3) *Citizenship Studies* 337

Grewal, Inderpal, 'Outsourcing Patriarchy' (2013) 15(1) *International Feminist Journal of Politics* 1

Jackson, Nicky Ali (ed), *Encyclopedia of Domestic Violence* (Taylor & Francis, 2007) doi.org/10.4324/9780203942215

Johnson, Holly, Natalia Ollus and Sami Nevala, *Violence Against Women: An International Perspective* (Springer, 2007)

Khan, Rao Arif Ali, *Honour Killing: Roots and Remedies* (Mittal Publications, 2012)

Mandhane, Renu, 'The Use of Human Rights Discourse to Secure Women's Interests: Critical Analysis of the Implications' (2004) 10 *Michigan Journal of Gender & Law* 275

Meetoo, Veena and Heidi Safia Mirza, '"There is Nothing 'Honourable' About Honour Killings": Gender, Violence and the Limits of Multiculturalism' (2007) 30 *Women's Studies International Forum* 187

Meyersfeld, Bonita, *Domestic Violence and International Law* (Bloomsbury Publishing, 2010)

Mohanty, Chandra Talpade, 'Under Western Eyes: Feminist Scholarship and Colonial Discourses' in Chandra Talpade Mohanty, Ann Russo and Lourdes Torres (eds), *Third World Women and the Politics of Feminism* (Indiana University Press, 1991) 51

Morgaine, Karen, 'Domestic Violence and Human Rights: Local Challenges to a Universal Framework' (2007) 34 *Journal of Sociology & Social Welfare* 109

Mutua, Makau, 'Savages, Victims, and Saviours: The Metaphor of Human Rights' (2001) 42(1) *Harvard International Law Journal* 201

Ray, Raka, 'Postcoloniality and the Sociology of Gender' in James W Messerschmidt et al (eds), *Gender Reckonings: New Social Theory and Research* (New York University Press, 2018) 73

Reddy, Rupa, 'Domestic Violence or Cultural Tradition? Approaches to "Honour Killing" as Species and Subspecies in English Legal Practice' in Aisha K Gill, Carolyn Strange and Karl Roberts (eds), *'Honour' Killing and Violence: Theory, Policy and Practice* (Palgrave Macmillan, 2014) 27

Soohoo, Cynthia, Catherine Albisa and Martha F Davis (eds), *Bringing Human Rights Home: A History of Human Rights in the United States* (University of Pennsylvania Press, 2008)

Spivak, Gayatri Chakravorty, 'Can the Subaltern Speak?' in C Nelson and L Grossberg (eds), *Marxism and the Interpretation of Culture* (University of Illinois Press, 1988) 271, doi.org/10.1007/978-1-349-19059-1_20

Ulrich, Jennifer L, 'Confronting Gender-Based Violence with International Instruments: Is a Solution to the Pandemic Within Reach?' (2000) 7(2) *Indiana Journal of Global Legal Studies* 629

Volpp, Leti, 'On Culture, Difference and Domestic Violence' (2002) 2(11) *Journal of Gender, Social Policy & the Law* 393

Watts, Charlotte and Cathy Zimmerman, 'Violence Against Women: Global Scope and Magnitude' (2002) 359(9313) *Lancet* 1232

Weissman, Deborah M, 'The Human Rights Dilemma: Rethinking the Humanitarian Project' (2004) 35 *Columbia Human Rights Law Review* 259

Wendt, Sarah and Lana Zannettino, *Domestic Violence in Diverse Contexts* (Routledge, 2015)

Cases

Lenahan (Gonzales) v United States of America (Inter-American Commission on Human Rights, Case No 12.626, 2011)

Treaties

Convention on the Elimination of All Forms of Discrimination against Women, opened for signature 18 December 1979, 1249 UNTS 13 (entered into force 3 September 1981)

Other

Charlesworth, Hilary, 'Are Women's Rights Human Rights?: International Law of Human Rights' (Speech delivered at the Alicia Johnson memorial lecture, Darwin, 3 September 1993)

Declaration on the Elimination of Violence against Women, GA Res 48/104, UN GAOR, 48th sess, 85th plen mtg, UN Doc A/RES/48/104 (20 December 1993)

Manjoo, Rashida, *Report of the Special Rapporteur on Violence against Women, its causes and consequences*, UN Doc A/HRC/26/38/Add.1 (1 April 2014)

National Coalition Against Domestic Violence, *Statistics* (2015) <ncadv.org/statistics>

National Domestic Violence Hotline, *Get the Facts & Figures* (2018) <www.thehotline.org/resources/statistics/>

Office of the High Commissioner for Human Rights, *Ratification Status for CEDAW* <tbinternet.ohchr.org/_layouts/TreatyBodyExternal/Treaty.aspx?Treaty=CEDAW&Lang=en>

UN Committee on the Elimination of Discrimination Against Women, *CEDAW General Recommendations Nos. 19 and 20*, 11th sess, UN Doc A/47/38 (24 June 1992)

UN Committee on the Elimination of Discrimination against Women, *Concluding Observations on the Combined Fourth and Fifth Periodic Reports of India*, UN Doc CEDAW/C/IND/CO/4-5 (24 July 2014)

United Nations Population Fund, *Gender-Based Violence* (2018) <www.unfpa.org/gender-based-violence>

United Nations Secretary-General, *In-Depth Study on All Forms of Violence against Women*, UN Doc A/61/122/Add.1 (6 July 2006)

United Nations Secretary-General, *Message on the International Day for the Elimination of Violence Against Women* (25 November 2007) <www.un.org/events/women/violence/2007/sg.shtml> (site discontinued)

Prosecuting elderly genocidaires: Why we should not reward the evasion of criminal justice

REBECCA KRIESLER

Abstract

In recent years, individuals in their 80s and 90s have been put before courts for their role in the Holocaust. Some now argue such legal proceedings are a waste of time, as the punishment available cannot match the crimes of genocide that are alleged, that the accused are too old to imprison. This paper explores the role of the trials of elderly genocidaires in contemporary society and whether it is just to punish these perpetrators. This paper examines the traditional theoretical framework applied to criminal punishment (protection, rehabilitation, deterrence and retribution), and whether or not these pillars remain relevant. The more modern theories of restorative justice and catharsis and the role of victims are then considered. This paper argues that these trials should continue to occur due to the important role they play in general deterrence of war crimes, punitive theories of retribution and the restorative benefit they have for survivors and their families.

Introduction

In recent years, individuals in their 80s and 90s have been put before courts for their role in the Holocaust. Some now argue such legal proceedings are a waste of time, as the punishment available cannot match the crimes of genocide that are alleged; that the accused are too old to imprison (Smeulers, 2008; Porcella, 2007). On the other hand, refusing to try elderly genocidaires may be seen as rewarding them for the continual evasion of the justice system as their culpability remains unchanged (Zuroff, 2012). The traditional theoretical approach to justice, generally, revolves around four pillars of punishment: protection, rehabilitation, deterrence and retribution (Daly, 2016). The three distinct, non-contingent rationales for continuing to punish genocidaires, in particular, are: the principled, non-consequentialist approach of

retribution (that crimes should be punished) (Wood, 2010); the utilitarian rationale of general deterrence (Dresser, 2001); and the restorative approach, looking at the benefit to victims (Barton, 2003).

This paper takes the position that it is important to continue prosecuting these perpetrators so that they face consequences for their crimes and victims are finally able to obtain justice. I address criticisms of the genocidaire trials by examining the aims of the criminal justice system that are no longer deemed relevant. Some people oppose holding trials, as protection, rehabilitation and personal deterrence are inconsequential given the age of the perpetrators. I consider, though, how the roles of retribution and general deterrence remain pertinent in order to raise awareness and increase accountability. I also examine emerging restorative theories of justice, whereby trials can empower victims by providing a platform to speak, and punishment can help them achieve closure. This paper argues that these trials retain their importance and perpetrators should continue to be prosecuted because the three key principles of justice (retribution, deterrence and restoration) remain pertinent regardless of time delays and the age of the accused. I primarily rely on case studies of the Holocaust and Cambodian Genocide, while also making reference to the Rwandan genocide.

Punishment

For punishment to be justifiable, it must provide some social utility. The advanced age and health conditions of the accused means that several relevant sentencing considerations are not applicable. Research has found there is an almost non-existent risk of the sick and elderly reoffending (Friedgood v NY State Bd, 2007). As such, they are predominantly incapable of posing a threat to society (Lundstrom, 1994). The role of protection or personal deterrence is therefore irrelevant, as there is no future risk to the community (Porcella, 2007).

Consequently, arguments about public safety do not factor into sentencing considerations. Rehabilitation is irrelevant as the accused's age means that they are unlikely to be reintegrated into society (Dressler, 2001). While these aims may not be satisfied by trying elderly genocidaires, for punishment to be warranted and just, it need not meet all the aims (Daly, 2016). Moreover, no justification in isolation is sufficient to establish that there is still an imperative to punish elderly genocidaires. The first justification of these trials is that since perpetrators of genocide have committed crimes, their actions warrant punishment in order to provide retribution.

Retribution

The principle of retribution is 'guilt deserves punishment for the sake of justice' (Sadurski, 1985, 223). Retributivism is a theory of justice that seeks to punish offenders because they *deserve* to be punished (Walker, 1980, 25–6). This pillar of justice takes a non-consequentialist approach to sentencing, as punishment is awarded due to its own innate worth, rather than for any consequential good or benefit it is claimed will result (Wood, 2010). Punishment is seen to restore a moral balance through the repaying of a debt (Dressler, 2001). Given this theory revolves around community morality, it is not time-sensitive, and punishment can be applied so long as the action is deemed wrong. Genocide is still a crime and thus culpability is not mitigated by the passage of time.

The trial of genocidaires is critical to denounce criminal conduct and articulate the international community's condemnation. For international law, punishment conveys 'the indignation of humanity for the serious violation of international humanitarian law' (Prosecutor v Miodrag Jokic, 2004). Perpetrators of genocide have committed a crime by attempting to cause the death of an entire culture, wherein their actions are in violation of humanity's ethical code. As Lord Denning stated: '[t]he ultimate justification of any punishment is the emphatic denunciation by the community of a crime' (cited in Cottingham, 1979, 245). Age does not impact culpability (Pertierra, 1995). Zuroff emphasises that the passage of time does not diminish Nazi criminals' guilt, 'they are just as guilty today as the day they committed their crime – and they do not deserve a prize for eluding justice for so long' (2012). Not implementing punishment means that they are not held to account for the heinous crimes they have committed.

Retributive aims of justice are reflected in ongoing prosecutions of the Khmer Rouge. The Khmer Rouge controlled Cambodia between April 1975 and January 1979 in a genocidal regime where approximately 1.7 million people were executed or died from hunger, exhaustion or disease (Meyer, 2008). The Extraordinary Chambers in the Courts of Cambodia (ECCC) was established in 2006, delayed almost 30 years after the Khmer Rouge lost power (Menzel, 2007). The Court was responsible for prosecuting five of the main perpetrators of the genocide, including Pol Pot's chief ideologist, and former president, Nuon Chea (Meyer, 2008). Although much time had lapsed, survivors have been at the heart of pushing for a trial. According to Menzel:

> a court trial is a clear message that Cambodian society does not want to simply 'reconcile' with the Khmer Rouge, but that it wants perpetrators to be called to account and punished according to their crimes. (Menzel, 2007, 225)

It is apparent that Cambodia is incapable of restorative justice or moving forward without some degree of retribution for perpetrators (Kek Galabru, 2000).

These trials are important as they allow victims to feel vindication and even some level of forgiveness after those responsible have been punished. Retributive theories of justice highlight that actions that violate the norms and ethics of society must be punished as a signal that perpetrators have done something wrong. It goes some way to right that balance so that the world may move on (Pertierra, 1995). Retribution is also seen as a prerequisite for forgiveness. This is evident in the Rwandan tribunal where one aim was to ensure perpetrators receive their 'just desserts' (Maogoto, 2004, 211). These trials' ultimate goal is to express moral condemnation, both domestically and by the international community, and to impose punishment on those who perpetuated the genocide (Mutua, 2000).

A criticism of retribution is the idea of proportionality; however, this is not a justification to fail to prosecute and thus create a more disproportional response. Central to retributive theories is the idea that perpetrators receive a sentence that is just in relation to the 'gravity of the crime and blameworthiness of the individual' (Goh, 2013, 45). However, Smeulers asks whether a sentence for genocide can ever be just (2008, 984). He poses the rhetorical question that, if one murder can result in the punishment of life imprisonment, how long should a sentence for complicity in hundreds of murders be? (Smeulers, 2008). Drumbl (2007) also argues that it is impossible to understand the sheer immensity and brutal nature of the crimes as they are incomprehensible. If you try to attach a numerical figure to these deaths it becomes 'unintelligible and immeasurable' (Drumbl, 2007, 156).

Even if the punishment is insufficient, it is still preferable to allowing perpetrators to escape reprimand completely. In relation to the Holocaust, Hannah Arendt wrote: 'for these crimes, no punishment is severe enough … That is, this guilt, in contrast to all criminal guilt, oversteps and shatters any and all legal systems' (cited in Bass, 2000, 13). However, the logic that no punishment is equivalent to the crime is not an argument for then allowing perpetrators to go unpunished. According to Ranki (1997, 32), the necessity of prosecuting is that choosing *not* to implicitly denotes an acceptance of the actions. It sends the message that there is nothing to prosecute; that the conduct is not deemed a crime (Ranki, 1997, 32).

In line with retributivism, neither the time delay nor condition or age of the accused is sufficient justification to refrain from prosecution. By virtue of having committed a crime, perpetrators deserve punishment. This view is echoed by Arendt (1963, 254), who believed that even if a punishment is insufficient, it is still important, as 'evil' actions destroy the natural balance, and it can only be restored through retribution. Society owes a duty to the moral order to rectify the imbalance and punish the criminal. The other duty society owes is to deter crimes and prevent them from transpiring.

Deterrence

If the trial and punishment of elderly genocidaires can discourage further atrocities then, under a utilitarian framework, they have merit and should continue. Deterrence means that the threat of punishments can dissuade potential perpetrators from committing crimes (Hassan, 1983). The object of deterrence is to maintain social control (Hassan, 1983). Utilitarianism holds that punishment is desirable only if it benefits the public (Dressler, 2001). Therefore if punishment can help prevent crimes it is justified.

There are two types of deterrence, special and general. Special deterrence is the punishment of an individual to prevent them reoffending (Smeulers, 2008). Special deterrence is not a relevant consideration for genocidaires as the 'likelihood of persons convicted here ever again being faced with an opportunity to commit war crimes … is so remote' (Prosecutor v Dragoljub Kunarac, 2001). Furthermore, the age of the accused means they are unlikely to reoffend. However, general deterrence, the idea of preventing others committing crimes, remains pertinent (Smeulers, 2008). This holds that when potential offenders see others being punished, they are dissuaded from committing a similar crime. This is summarised by Bentham: 'Punishment must be the object of dread more than the offense is an object of desire' (1951, 332).

Deterrence was a major goal of post–First World War international criminal law, aiming to prevent people from committing future war crimes. Deterrence is especially relevant for international crimes, as there is a perception that international law is not enforceable and therefore actions will go unpunished (Hassan, 1983). Current trials are particularly effective deterrents, as they demonstrate that there is no impunity for offenders, regardless of age or delay (Carmichael, 2011). During the trial of the Auschwitz guards, Reinhold Hanning, a prosecution lawyer, commented that the trial was an opportunity to reaffirm the horror and prevent its repetition: 'throwing light on what happened, with ensuring that something like this never happens again', (Goldbach cited in Connolly, 2016). The idea that you could be punished at any moment, even after you think you have escaped punishment, is a powerful deterrent. The hope is that these trials will force people to realise their own security and freedom will be in jeopardy by punishment and this will prevent future crimes against humanity occurring (Schense and Carter, 2016).

The trial of the leadership of the Khmer Rouge shows the role of deterrence as a sentencing consideration. The Joint Statement of the Prosecutors of the ICC,[1] the ICTY,[2] ICTR[3] and SCSL[4] in 2004, stated:

> People of the world are entitled to a system that will deter grave international crimes and hold to account those who bear the greatest responsibility. Only when a culture of accountability has replaced the culture of impunity can the diverse people of the world live and prosper together in peace (Joint Statement, 2004).

This statement is exemplified in the trial of Nuon Chea, Ieng Sary, Khieu Samphan and Hun Sen, four senior members of the Khmer Rouge, who were responsible for the Cambodian genocidal regime (Beachler, 2014). Although the defendants were elderly and frail, the court case was deemed important to combat the myth of Khmer non-liability pervading Cambodia (McCargo, 2011). According to McCargo, 'the hope was that the imposition of legal processes and punishment against key Khmer Rouge cadre would undermine the culture of impunity in Cambodia' (McCargo, 2011, 613–27). It was important to show that even the heads of regimes could be held accountable, and that everyone was subject to the rule of law. Neier notes the importance of trials like this in ensuring social cohesion, as they ensure the basic rules cannot be disregarded without consequences (1998). It is crucial that citizens respect the criminal justice system and trust in its efficiency in order for a nation to rebuild and to ensure stability.

The fear of social punishment acts to stop people committing crimes. From a utilitarian framework, these trials are warranted as they can prevent future deaths. This is also seen in Rwanda where punishments aimed to be harsh enough to 'convince [the Rwandan people] that they have too much to lose if they succumb to illegal activities' (International Centre for the Study and the Promotion of Human Rights and Information, 1997, 42). Trials are an effective mechanism for deterring crime as they hold genocidaires accountable and thus discourage others. Furthermore, these trials deter the denial of genocides.

Holocaust denial is another reason why deterrence is critical, because such denial threatens to delegitimise the experiences of victims, increases anti-Semitism and may ultimately pave the way for future atrocities. There has been a growing number of people who deny the Holocaust, claiming it was fabricated and no crime had been committed (Ranki, 1997). The ongoing prosecution of Nazis can be justified as it helps

[1] International Criminal Court.

[2] International Criminal Tribunal for the Former Yugoslavia.

[3] International Criminal Tribunal for Rwanda.

[4] Special Court for Sierra Leone.

combat this threat. Ranki stresses that in a context of growing denialism, reaffirming the identity of the perpetrators and victims has important civic value (1997, 31). Trials are crucial in terms of gathering evidence and placing statements on an official historical record. One of Nuremberg's central successes was to prevent and counter denial of the Holocaust because of its meticulous collection of evidence (Goldstone, quoted in Gearty 2007). Ongoing cases are important for a similar reason, in providing concrete evidence of the Holocaust and ensuring it remains at the centre of attention. Irwin Cotler believes that every time a Nazi war criminal is convicted we repudiate the denial movement. However the inverse is also true, every time society opts not to prosecute, inferences are drawn that no crime has been committed (Cotler, 1988, 70). These trials must continue, as the prosecution of Nazis now is important to rebuff modern denial attempts and ensure victims' experiences are not erased. Therefore, these trials have utility as they prevent denial of genocide and potentially prevent future genocides.

Empowering victims

Finally, genocidaires should be subject to trial and punishment for the benefit of survivors and their families in terms of closure and empowerment. Bandes looks at how the theory of closure has been embraced as a goal for victims and survivors that the legal system ought strive towards (Bandes, 2009, 2). According to restorative justice philosophy, 'closure and emotional healing' is vital for restorative outcomes (Barton, 2000). The lack of resolution or convictions for victims of genocide can have a debilitating effect on their wellbeing or ability to process the trauma and move on (Maogoto, 2004, 211). Perpetrators never being caught or brought to justice can prevent catharsis. Punitive justice can help individuals move on, can assuage anger and help victims reach closure.

The restorative impact these trials has on survivors and their families can be demonstrated in the trial of Auschwitz guard Reinhold Hanning in 2016. This case saw the conviction of a 94-year-old, former death camp guard as accessory to the murder of 170,000 people (Connolly, 2016). A survivor of the Holocaust, Hedy Bohm, flew from Canada to testify at the trial and told reporters she was 'grateful and pleased by this justice finally after 70 years' (cited in Connolly, 2016). The trial allowed people whose families had been killed to reach some form of peace, knowing perpetrators were being punished. Equally as important, the trial also gave survivors the opportunity to speak out by testifying and talking to the media (Connolly, 2016). These trials are an official acknowledgement and recognition of wrongdoing, 'which for many victims is the beginning of their healing process' Goldstone, quoted in Gearty 2007, 3). It is a powerful acknowledgement that the state and the community are on the victims' side, by recognition of the events that transpired and the resultant harms.

Empowering victims works to undermine the voices of deniers and provides a sense of support within the community, rather than isolation. According to special prosecutor Schrimm, 'one owes it to the survivors and the victims not to simply say "a certain time has passed, it should be swept under the carpet"' (cited in Nagorski, 2016, 316). For victims who have had to deal with the repercussion of these atrocities for decades, the time delay does not mitigate the guilt or need for justice. If anything it exacerbates it, as they are confronted with their culprits' freedom and lack of consequences.

Although trials are criticised as being ineffective mechanisms for catharsis, this is undermined by the central role victims play in pushing for these cases to be heard. According to Murphy, is that trials are a 'poor vehicle for authentic expression of emotion' (2012, 157). Murphy believes that victims place too much importance on trials to help them reach closure, as this can leave them feeling empty and aimless (2012). It is fair to say that for many people a trial is insufficient, as it cannot reverse the losses suffered. However, if trials can provide victims with any form of closure or assuage any suffering, they are justified regardless of whether they fall short of remedying the entire situation. As Goldstone writes, 'the main customers of war crimes trials, like any trials, are the victims' (Goldstone, quoted in Gearty 2007, 3). Given victims are often at the forefront of campaigns for punishments and trials, it seems evident that they do provide some benefit (Menzel, 2007).

A further criticism is the traumatic effects trials can have on survivors who are forced to relive terrible experiences; however, the opt-in nature of these court cases provides relief to this criticism. It is said that these trials can re-traumatise victims, especially through the distressing effect of testifying (Brehm, Uggen & Gasanabo, 2014). However, these systems have optional participation, which means victims are not forced to relive horrors unless they chose to participate. Many victims say they feel empowered when confronting perpetrators (Catani, 2012). In the trial of Groning, Kathleen Zahavi was able to directly address the man who was responsible for her parents' death:

> I hope the images of what went on there will stay with you for the rest of your days. You were allowed in your freedom to grow old. My parents weren't allowed that. (cited in Withnall, 2015)

To have the opportunity to express your thoughts and to personally hold perpetrators to account can be invaluable to victims and is an opportunity only granted through trials.

The ECCC is a pertinent example of a court being used to empower victims. The ECCC was distinct from most other genocide courts as it focused on victims and included a victim participation scheme (Boyle, 2006). The ECCC enacted victim participation provisions, permitting survivors to file complaints, apply to

become witnesses or become a civil party.[5] In one case, over 4,000 victims sought civil party status (ECC, Press Release, 24 June 2011). Allowing victims to participate as civil parties is unprecedented in criminal justice and provided a unique opportunity for victims to have their voice heard (Nguyen & Sperfeldt, 2014). Karstedt called this new scheme 'a road from absence to presence, and from invisibility to the visibility of victims' (2010, 17). The ECCC allowed victims to express and process traumatic experiences and facilitated some survivors achieving closure (Nguyen & Sperfeldt, 2014). Furthermore, they had the opportunity to present their personal experiences and memories in a public forum, thus being able to express and validate their trauma (Nguyen & Sperfeldt, 2014). Allowing different minorities to present the distinct nature of the harm they suffered under the Khmers was important as it meant their personal experiences were acknowledged rather than homogenised (Mohan, 2008). Given the trial was decades after the event, the trial was an opportunity to revisit past events and reflect and meant that victims were able to move on and feel that a sense of justice had been achieved (Cataoni, 2012). These trials are a unique occasion for victims to reach closure, allowing them to voice their experiences and see their perpetrators punished.

Conclusion

Genocidaire trials should continue to occur due to the important role they play in general deterrence of war crimes, punitive theories of retribution and the restorative benefit they have for survivors and their families. These elements are not impacted by the passage of time. Although punishing elderly genocidaires is controversial as they no longer pose a threat to society, these trials still provide significant benefits to victims and the community. Retribution is morally just as trials allow mass murderers to receive punishment.

The sentencing of genocidaires help to prevent future crimes, as the accountability of perpetrators deters actions and the public nature of these trials combats denialism. These trials help victims achieve a sense of justice through having their experiences acknowledged and seeing their perpetrators held to account. Moreover, the principals of deterrence, retribution and restoration are equally applicable in relation to other trials that involve time delays or elderly offenders. In the prosecution of war crimes, historical sexual assaults, murder cases that remain unsolved and other indicatable offences, it is important that cases are

[5] Extraordinary Chambers in the Courts of Cambodia, Internal Rules (Ver 4) (adopted 11 September 2009).

still brought regardless of delays. Time should not prove a barrier to bringing legal action when victims and their families continue to suffer for the acts of violence perpetrated against them.

Bibliography

Arendt, H. (1963). *Eichmann in Jerusalem: A Report on the Banality of Evil*. Viking Press.

Bandes, S. (2009). Victims, 'Closure,' and the Sociology of Emotion. *Law and Contemporary problems,* 72(2), 1–26.

Barton, S. (2003). *Restorative Justice – The Empowerment Model.* Hawkins Press.

Bass, G. (2000). *Stay the Hand of Vengeance: The Politics of War Crimes Tribunals*. Princeton University Press.

Beachler, D. (2014). The Quest for Justice in Cambodia: Power, Politics and the Khmer Rouge Tribunal. *Genocide Studies and Prevention: An International Journal,* 8(2), 67–80. doi.org/10.5038/1911-9933.8.2.7.

Bentham, J. (1951). 'Theory of Legislation' in Cohen, M & Cohen, F. (eds), *Reading In Jurisprudence and Legal Philosophy*. Cohen & Cohen.

Boyle, D. (2006). The Rights of Victims: Participation, Representation, Protection, Reparation. *Journal of International Criminal Justice,* 4. doi.org/10.1093/jicj/mql006.

Brehm, HN, Uggen, C, & Gasanabo, J-D. (2014). Genocide, Justice, and Rwanda's Gacaca Courts. *Journal of Contemporary Criminal Justice,* 30(3), 333–52. doi.org/10.1177/1043986214536660.

Carmichael, R. (2011, November 21). Khmer Rouge Trial Opens: Key Charges Might Never be Heard. *Voice of America.* Retrieved from www.voacambodia.com/a/cambodias-khmer-rouge-tribunal-legacy-found-in-communities-around-country-101878938/1357107.html.

Catani, L. (2012). Victims at the International Criminal Court. Some Lessons Learned from the Lubanga Case. *Journal of International Criminal Justice,* 10, 905. doi.org/10.1093/jicj/mqs053.

Charles Friedgood v New York State Board of Parole, NY Slip Op 08158 [45 AD3d 947] (2007).

Connolly, K. (2016, June 17). Auschwitz Guard Jailed for Five Years in Holocaust Murder Trial. *The Guardian*. Retrieved from www.theguardian.com/world/2016/jun/17/auschwitz-guard-reinhold-hanning-jailed-holocaust-auschwitz.

Cotler, I. (1988). Symposium – Holocaust and Human Rights Law: The First International Conference. *Boston College Third World Journal*, 8, 68.

Cottingham, J. (1979). The Varieties of Retribution. *Philosophical Quarterly*, 29, 238–46. doi.org/10.2307/2218820.

Daly, K. (2016). *Criminal Justice System: Aims and Processes*. Lawbook Co.

Dressler, J. (2001). *Understanding Criminal Law.* Lexis Law Pub (3rd Ed).

Drumbl, M. (2007). *Atrocity, Punishment and International Law.* Cambridge University Press.

ECCC. (Press Release, 24 June 2011). Pre-Trial Chamber Overturns Previous Rejection of 98 Percent of Appealing Civil Party Applicants in Case 002.

Findlay, M. (2009). Activating a Victim Constituency in International Criminal Justice. *International Journal of Transitional Justice,* 3, 183. doi.org/10.1093/ijtj/ijp008.

Gearty, C (Presenter). (2007, May 3). Interview with Richard Goldstone, Leslie Vinjamuri & Anthony Dworkin, *Do War Crime Trials Do More Harm than Good?* London School of Economics.

Goh, J. (2013). Proportionality – An Unattainable Ideal. *Manchester Student Law Review,* 2, 41-72.

Hassan, F. (1983). The Theoretical Basis of Punishment in International Criminal Law. *Case Western Reserve Journal of International Law*, 15(1), 39-60.

International Centre for the Study and the Promotion of Human Rights and Information. (1997). *The Genocide and the Crimes against Humanity in Rwandan Law*. ASSEPCA.

Karstedt, S. (2017). From Absence to Presence, From Silence to Voice: Victims in International and Transitional Justice since the Nuremberg Trials. *International Review of Victimology,* 17, 9. doi.org/10.1177/026975801001700102.

Kek Galabru. (2000). President, LICADHO (Cambodian League for the Promotion and Defense of Human Rights), as quoted in 'Cambodians talk about the Khmer Rouge trial', *Phnom Penh Post*, 4–17 February 2000, 12.

Lundstrom, S. (1994). Dying To Get Out: A Study on the Necessity, Importance, and Effectiveness of Prison Early Release Programs for Elderly Inmates and Inmates Suffering from HIV Disease and Other Terminal-Centered Illnesses. *Brigham Young University Journal of Public Law,* 9(1), 155.

Maogoto, J. (2004). The International Criminal Tribunal for Rwanda: A Paper Umbrella in the Rain? Initial Pitfalls and Brighter Prospects. *Nordic Journal of International Law* 73(2). doi.org/10.1163/1571810042321177

McCargo, D. (2011). Politics by Other Means? The Virtual Trials of the Khmer Rouge Tribunal. *International Affairs*, 87(3), 613–27. doi.org/10.1111/j.1468-2346.2011.00993.x.

Menzel, J. (2007). Justice Delayed or Too Late for Justice? The Khmer Rouge Tribunal and the Cambodian 'Genocide' 1975–79. *Journal of Genocide Research*, 9(2). doi.org/10.1080/14623520701368602.

Meyer, W. (2008). The Khmer Rouge Tribunal – Cambodia in the Shadow of History. *International Reports* 3.

Mohan, M. (2008). Re-Constituting the 'Un-Person': The Khmer Krom and the Khmer Rouge Tribunal. *Singapore Yearbook of International Law,* 12, 43.

Murphy, J. (2012). *Punishment and the Moral Emotions: Essays in Law, Morality and Religion*. Oxford University Press. doi.org/10.1093/acprof:osobl/9780199764396.001.0001.

Mutua, M. (2000). From Nuremberg to the Rwanda Tribunal: Justice or Retribution? *Buffalo Human Rights Law Review* 6(77).

Nagorski, A. (2016). *In Pursuit: The Men and Women Who Hunted the Nazis*. Simon & Schuster.

Neier, A. (1998). *War Crimes: Brutality, Genocide, Terror, and the Struggle for Justice*. Times Books.

Nguyen, L, & Sperfelt, C. (2014). Victim Participation and Minorities in Internationalized criminal Trials: Ethnic Vietnamese Civil Parties at the Extraordinary Chambers in the Courts of Cambodia. *Macquarie Law Journal,* 14.

Pertierra, C. (1995). Do the Crime, Do the Time: Should Elderly Criminals Receive Proportionate Sentences? *Nova Law Review* 19, 793.

Porcella, K. (2007). The Past Coming Back to Haunt Them: The Prosecution and Sentencing of Once Deadly but Now Elderly Criminals. *St. John's Law Review,* 81, 369.

Prosecutor v Dragoljub Kunarac, Radomir Kovac and Zoran Vukovic (Trial Judgment). International Criminal Tribunal for the former Yugoslavia (ICTY). (22 February 2001).

Prosecutor v Miodrag Jokic (Sentencing Judgement). International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of Former Yugoslavia. (18 March 2004).

Ranki, V. (1997). Holocaust History and The Law: Recent Trials Emerging Theories. *Law & Literature*, 9(1), 15–44. doi.org/10.1080/1535685X.1997.11015798.

Sadurski, W. (1985). *Giving Desert Its Due*. Dordrecht: Reidel.

Schense, J, & Carter, L. (2016). *Two Steps Forward, One Step Back: the Deterrence Effect of International Criminal Tribunals*. International Nuremberg Principles Academy.

Smeulers, A. (2008). Punishing the Enemies of All Mankind. *Leiden Journal of International Law*, 21, 971–93. doi.org/10.1017/S0922156508005487.

Walker, N. (1980). *Punishment, Danger and Stigma: The Morality of Criminal Justice*. Rowman & Littlefield.

Withnall, A. (2015, May 13). Oskar Groning Trial: 'Bookkeeper of Aushwitz' Reduced to Tears by Dramatic Testimony of Death Camp Survivor. *The Independent.* Retrieved from www.independent.co.uk/news/world/europe/oskar-groning-trial-bookkeeper-of-auschwitz-reduced-to-tears-by-dramatic-testimony-of-death-camp-10247586.html.

Wood, D. (2010). Punishment: Consequentialism. *Philosophy Compass,* 5(6). doi.org/10.1111/j.1747-9991.2010.00287.x.

Zuroff, E. (2012, September 25). Why Nazi Hunting is Still a Worthy Pursuit. *CNN International.* Retrieved from edition.cnn.com/2012/09/24/opinion/zuroff-nazi-hunting/index.html.

A sugar tax can answer Australia's obesity problem

CLARE LANGLEY

Abstract

This article aims to review literature examining the impact of a sugar-sweetened beverage (SSB) tax in order to make recommendations to the Australian Government about the efficacy of a sugar tax in Australia. Due to Australia's ageing population and the changed lifestyle of the population in the twenty-first century, chronic diseases have become the largest cause of death in Australia. A significant common risk factor in the four major chronic diseases is obesity. Obesity is caused by a number of factors, some of which are preventable, including nutrition. Thus, strategies that aim to prevent obesity have become a significant area of chronic disease prevention. One such strategy is the introduction of a sugar tax. This article argues that SSB taxes have a positive impact on lowering obesity. A 20 per cent tax on SSBs is the recommendation to the Australian Government, with revenue being used to fund other obesity prevention strategies. Although the government has previously quashed a proposed sugar tax due to belief that the tax would have little positive impact on Australian population, this article indicates that the tax would provide an inexpensive and effective method of addressing obesity prevalence.

Introduction

The largest public health issue for Australia today is the prevention and treatment of chronic disease (AIHW, 2014). A common risk factor between the four major chronic diseases in Australia (cardiovascular disease, cancer, chronic obstructive pulmonary disease and diabetes) is obesity (AIHW, 2014). The prevalence of obesity has dramatically increased in the past decade; most alarming is the increase of childhood obesity (Australian Bureau of Statistics, 2015). Despite this, there has been little government intervention involved in the primary prevention of the disease (Australian Government, 2013). A sugar-sweetened beverage (SSB) tax, dubbed a 'sugar tax', is a proposed method for reducing obesity by using a taxation model (Sarlio-Lähteenkorva & Winkler, 2015).

This article aims to determine if implementing a sugar tax in Australia would improve public health outcomes, by evaluating research from a number of countries with sugar taxes. The success of the tax will be evaluated on two fronts: the effect on the prevalence of obesity; and the economic impact. All literature discussed is peer reviewed, unbiased and from the last decade. Review of this evidence points

to a recommendation that a 20 per cent sugar tax should be implemented in Australia. This would reduce the prevalence of obesity and create revenue which can be used to create other policies addressing obesity.

Reasons for a sugar tax

Many public health experts and health care researchers have determined that SSBs have had a significant impact on the increase in obesity in Western countries over the last decade (Andreyeva, Chaloupka & Brownell, 2011). The theory highlights the correlation between increase in the amount of SSBs consumed and the increased prevalence of obesity (Pereira, 2014). However, overall there has been an increase in all high-sugar processed foods, not just SSBs (Andreyeva et al., 2011). The major reasons SSBs have become the main target are their high added sugar content and low satiety, combined with their high energy content (Malik, Popkin, Bray, Despres & Hu, 2010). These factors, along with their liquid form, create more rapid absorption of carbohydrates compared to other high-sugar products (Malik et al., 2010). Not only do these characteristics of SSBs contribute to weight gain, they also increase risk factors of type 2 diabetes and cardiovascular disease (Malik et al., 2010). Due to the growing evidence supporting this understanding, taxing SSBs becomes a compelling option to tackle obesity.

Taxation is an effective and inexpensive public health method which has been shown to alter food consumption (Andreyeva et al., 2011). Currently 26 countries have enacted some kind of sugar tax, the first one beginning in the early 1990s (Cabrera Escobar, Veerman, Tollman, Bertram & Hofman, 2013). This has allowed researchers to study whether the tax successfully alters behaviour and can reduce obesity rates (Cabrera Escobar et al., 2013). Results from such studies led the World Health Organization, in their 2016 commission into ending childhood obesity, to recommend implementing ‘an effective tax on sugar-sweetened beverages’ (WHO, 2016).

Impact on obesity

Current literature on sugar taxes demonstrates that they have a moderate but considerable impact on obesity. The main aim of a sugar tax is to reduce obesity by reducing the number of SSBs consumed. A number of studies have aimed to determine if this works in reality. One such study by Fletcher, Frisvold & Tefft (2010) used a correlation study method to evaluate the impact of sugar tax in six US states. The study used national databases from six states with a 1 per cent sugar tax from 1990 to 2006 to correlate the amount of SSBs sold with average body mass index (BMI) and the percentage of people classed as obese and overweight, from a number of demographics including gender, age and income (Fletcher et al., 2010). Results from the study indicated a decrease in BMI of 0.003 percentage points, in obesity of 0.01 points and in people classed as overweight of 0.02 points (Fletcher et al., 2010). While the results demonstrate a moderate decrease in obesity, the tax implemented was not nationwide and

was only 1 per cent, which implies that a greater tax of 20 per cent – as proposed by the World Health Organization (2016) – would likely reduce the average BMI to a more significant degree (Fletcher et al., 2010).

In researching the impacts of a 20 per cent SSBs tax, a study in the UK by Briggs et al. (2013) found promising results. The research aimed to estimate the effect of 20 per cent sugar tax on the prevalence of overweight individuals in the UK prior to its implementation in early 2018 (Briggs et al., 2013). Researchers used modelling software to estimate the effect of the tax by using data from several national surveys recording purchases of SSBs and BMI (Briggs et al., 2013). The results indicate a predicted reduction of obesity by 1.3 per cent and a reduction of 104 mL of SSB consumption per person per week (Briggs et al., 2013). However, the concluding statement of the research warned that a sugar tax 'should not be seen as a panacea' (Briggs et al., 2013).

The sugar tax has also positively affected lower-income individuals, as seen through calculating the difference in prevalence of obesity in different income brackets (Briggs et al., 2013). This indicates an important consideration for the introduction of a sugar tax in Australia, where obesity effects individuals from low income brackets most significantly (Lal et al., 2017). In fact, low socio-economic status (SES) is a major risk factor for obesity (Blecher, 2015; Lal et al., 2017). Along with this, it is assumed that for low-income individuals money is the biggest predictor of behaviour (Blecher, 2015; Lal et al., 2017). However, the research found no significant difference between income groups (Briggs et al., 2013).

The most significant research into the relationship between sugar tax and the mitigation of socio-economic disparities in obesity in Australia is perhaps that conducted by Lal et al. (2017). These researchers used modelling techniques to estimate the health benefits of a 20 per cent sugar tax in Australia with particular focus on different SES groups (Lal et al., 2017). They used data from a number of Australia databases to determine baseline obesity rates and SSB consumption from different SES groups, and, using models created, determined the amount of half-adjusted life years (HALYs) gained. HALYs increase as BMI decreases, thus rendering it a measure of obesity prevalence (Lal et al., 2017). Results indicated a total gain of 175,300 HALYs, equating to 111,700 years of life saved (Lal et al., 2017). This was predicted to be most significant in individuals from low SES groups, and in both men and women (Lal et al., 2017). While research on the sugar tax has demonstrated it produces a decrease in the prevalence of obesity, for the tax to be considered viable, its economic impact must also be considered.

Impact on economy

The economic impact of a sugar tax is also significant: this includes the impact of tax on the consumer level, the effect on the health care system and the revenue produced. A major criticism of food taxes such as the sugar tax is that taxing a product does not actually result in changes to that product price (Sarlio-Lähteenkorva & Winkler, 2015). In relation to the sugar tax, it is often absorbed by SSB

companies and not seen in retail prices, or offset by increasing price of drinks which use sugar-free sweeteners (Sarlio-Lähteenkorva & Winkler, 2015). However, research by Alvarado et al. (2017) analysed price changes of SSBs and non-SSBs before and after a sugar tax of 10 per cent was implemented in Barbados. Using data of beverage prices from a major supermarket chain from 2014 to 2016, results showed that, before-tax, the SSBs and non-SSBs had similar prices; after the tax was implemented, there was initially a price increase of 3 per cent to SSBs and slight decrease to the price of non-SSBs (Alvarado et al., 2017). This trend continued for SSBs, reaching 5.9 per cent by the end of the study. while non-SSB prices remained stable after their initial drop (Alvarado et al., 2017). This demonstrates that sugar tax can successfully change retail prices (Alvarado et al., 2017). Another consumer consideration relating to the sugar tax is what customers purchase instead of SSBs (Singhal & Joshi, 2017). Research by Singhal & Joshi (2017) determined that while purchases of SSBs decreased, purchases of non-SSBs increased by 4 per cent in Mexico after a sugar tax was implemented: the majority of non-SSBs purchased was water (Singhal & Joshi, 2017).

The effect sugar tax would have on healthcare costs also deserves consideration. The previously reviewed Lal et al. (2017) study also modelled the effect of sugar tax on healthcare costs in Australia. In 2015, AUD 5.3 billion was spent on treating diseases caused by obesity in Australia (Lal et al., 2017). This price will only increase as the prevalence of obesity increases (Lal et al., 2017). The researchers developed a statistical modelling method called the CRE-Obesity model, which allowed them to predict an AUD 1.733 billion decrease in healthcare cost over the lifetime of the current population (Lal et al., 2017). While this is only a predicted outcome of healthcare costs, even if over-optimistic the results suggest a significant impact on the healthcare system. Since sugar taxation is recent in countries that have implemented it, there has been limited research effectively evaluating its impact on healthcare costs.

Another economic consideration is the national financial gain from the sugar tax, which could be used to create further programs aiming to reduce obesity. Within Australia, the revenue was estimated by Lal et al. (2017) to be AUD 642.9 million. This was estimated by taking into consideration administrative costs and changes in healthcare costs (Lal et al., 2017). Within Australia, the revenue from a sugar tax could be used to implement other programs to address obesity (Backholer, Blake & Vandevijvere, 2017). Backholer et al. (2017) reviewed revenue obtained from a number of countries and where that revenue was used. In the US states of Colorado and Washington, the revenue was used to fund programs involved in the prevention of obesity, targeting those most at risk and in programs educating children about healthy diets (Backholer et al., 2017). Andreyeva et al. (2011) determined possible uses for revenue such as using it to offset the price of more expensive, healthier foods, making them more affordable than cheap, unhealthy foods (Andreyeva et al., 2011).

Recommendations

The specific recommendations of this literature review are to implement a 20 per cent sugar tax to combat obesity, with revenue from this tax being used to address obesity. In 2013, in a report into obesity presented to the House of Representatives, one of the 20 recommendations to the government was a tax incentive program (Australian Government, 2013). The response from the (then) government was as follows:

> While obesity does involve significant health and productivity costs, the relationship between these costs and the consumption of particular products is complex. The risk of obesity is affected by lifestyle, such as diet and physical activity, as well as inherited and social influences. This makes it very difficult to estimate spillover costs, if any, of identifiable foods or food types. (Australian Government, 2013, p. 13)

Given the evidence outlined in this review, there is now sufficient evidence to indicate the contrary; that, while there are other risk factors involved in the cause of obesity, a tax on sugar will successfully reduce obesity prevalence in Australia (Cabrera Escobar et al., 2013). Along with this, a sugar tax will create significant revenue and reduce healthcare costs giving the government more money to implement other programs to address the other risk factors of obesity (Lal et al., 2017).

All research reviewed points to SSBs increasing obesity and that taxing such products will reduce obesity rates. There is, however disagreement regarding whether the sugar tax alone can tackle obesity (Briggs et al., 2013). There are currently a limited number of longitudinal studies, due to sugar taxes being a recent strategy, while statistical models used only predicted outcomes. Inaction on the issue of obesity in Australia is not an option, as obesity rates and related chronic diseases will continue to rise without government intervention. The sugar tax is an inexpensive, simple method which has been proven to effect prevalence of obesity and create significant revenue (AIHW, 2014; Andreyeva et al., 2011). A 20 per cent sugar tax should be enacted in Australia as soon as possible to reduce obesity rates, mostly significantly in low SES groups. Moreover, the tax will create revenue for further public health efforts to combat obesity.

References

AIHW. (2014). Chronic disease — Australia's biggest health challenge. *Australia's Health 2014*, *No. 14* (Cat no. AUS 178), 11. Retrieved from www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129547726.

Alvarado, M, Kostova, D, Suhrcke, M, Hambleton, I, Hassell, T, Samuels, TA, … Unwin, N. (2017). Trends in beverage prices following the introduction of a tax on sugar-sweetened beverages in Barbados. *Preventive Medicine*, *105*, 23–25. doi.org/10.1016/j.ypmed.2017.07.013.

Andreyeva, T, Chaloupka, FJ, & Brownell, KD. (2011). Estimating the potential of taxes on sugar-sweetened beverages to reduce consumption and generate revenue. *Preventive Medicine*, *52*(6), 413–16. doi.org/10.1016/j.ypmed.2011.03.013.

Australian Bureau of Statistics. (2015). *National health survey: First results. International statistical classification of diseases and related health problems*. doi.org/ABD Cat no. 4364.0.55.001.

Australian Government. (2013). *Australian Government response to the House of Representatives Standing Committee on Health and Ageing report: Weighing it up: Obesity in Australia*.

Backholer, K, Blake, M, & Vandevijvere, S. (2017). Sugar-sweetened beverage taxation: An update on the year that was 2017. *Public Health Nutrition*, *20*(18), 3219–24. doi.org/10.1017/S1368980017003329.

Blecher, E. (2015). Taxes on tobacco, alcohol and sugar sweetened beverages: Linkages and lessons learned. *Social Science and Medicine*, *136–137*, 175–79. doi.org/10.1016/j.socscimed.2015.05.022.

Briggs, ADM, Mytton, OT, Kehlbacher, A, Tiffin, R, Rayner, M, & Scarborough, P. (2013). Overall and income specific effect on prevalence of overweight and obesity of 20% sugar sweetened drink tax in UK: Econometric and comparative risk assessment modelling study. *BMJ (Online)*, *347*, 1–17. doi.org/10.1136/bmj.f6189.

Cabrera Escobar, MA, Veerman, JL, Tollman, SM, Bertram, MY, & Hofman, KJ. (2013). Evidence that a tax on sugar sweetened beverages reduces the obesity rate: A meta-analysis. *BMC Public Health*, *13*(1), 1–10. doi.org/10.1186/1471-2458-13-1072.

Fletcher, JM, Frisvold, D, & Tefft, N. (2010). Can soft drink taxes reduce population weight? *Contemporary Economic Policy*, *28*(1), 23–35. doi.org/10.1111/j.1465-7287.2009.00182.x.

Lal, A, Mantilla-Herrera, AM, Veerman, L, Backholer, K, Sacks, G, Moodie, M, … Peeters, A. (2017). Modelled health benefits of a sugar-sweetened beverage tax across different socioeconomic groups in Australia: A cost-effectiveness and equity analysis. *PLoS Medicine*, *14*(6), 1–17. doi.org/10.1371/journal.pmed.1002326.

Malik, VS, Popkin, BM, Bray, GA, Despres, J-P, & Hu, FB. (2010). Sugar-sweetened beverages, obesity, type 2 diabetes mellitus, and cardiovascular disease risk. *Circulation*, *121*(11), 1356–64. doi.org/10.1161/CIRCULATIONAHA.109.876185.

Pereira, MA. (2014). Sugar-sweetened and artificially-sweetened beverages in relation to obesity risk. *Advances in Nutrition*, *5*, 797–808. doi.org/10.3945/an.114.007062.

Sarlio-Lähteenkorva, S, & Winkler, JT. (2015). Could a sugar tax help combat obesity? *BMJ (Online)*, *351*, 8–10. doi.org/10.1136/bmj.h4047.

Singhal, A, & Joshi, S. (2017). Taxing sugary beverages reduces their purchase, especially among poor households. *Journal of Evidence-Based Dental Practice*, *17*(2), 145–47. doi.org/10.1016/j.jebdp.2017.04.002.

WHO. (2016). Report of the commission on ending childhood obesity. *World Health Organization*, 30.

Gold's influence on Australian economic development in the nineteenth century

TOM GOODWIN

Abstract

This article seeks to set out the significant ways in which the gold industry, particularly during the booms of the nineteenth century, affected Australia's development. I argue that the most substantial effects of the gold rush affected two broad aspects of Australian society. First, the gold rush impacted the Australian economy by bringing substantial wealth to Australian shores as an export, stimulating secondary industries, driving population growth and restructuring the manufacturing sector. Second, the gold rush helped shape Australia's socio-political climate. Compared with the previously hegemonic wool industry, gold provided better economic wages for the working class. These wage increases spread to other industries and, with increased population in Australia, a new era began wherein the working class expected high wages. I posit that this was due to the labour movement dominating Australia's political left. I suggest that the labour movement's influence on the political left has fostered more stable social outcomes in comparison with Argentina, an economy that developed a more radical political left.

Introduction

Edward Hargraves's discovery of gold on the fields of Ophir in 1851[1] flagged the beginning of arguably the most important period of Australian economic development. Economic historians widely agree that gold's impact on the Australian economy was as significant as any other factor since colonisation. The task that remains is to pinpoint exactly how gold shaped Australia's economy. I argue that gold's influence on the Australian economy was twofold. Gold not only brought unprecedented wealth to the colony's shores, but fundamentally restructured consumption, production and the political and social structure of Australia. In discussing how gold altered economic development in the nineteenth century, I will compare historical trends against a counterfactual scenario in which gold was not mined, ceteris paribus.

This essay identifies two broad themes of economic change instigated by the emergence of the gold industry, which caused a thirteenfold increase in the value of Australian exports, and a threefold population increase from 1851 to 1860.[2] First, I examine the direct effects of gold on Australia's pattern

[1] Sumner J La Croix, 'Property Rights and Institutional Change During Australia's Gold Rush'. *Explorations In Economic History* 29 (1992): 5.

[2] Rodney Maddock and Ian McLean. 'Supply-Side Shocks: The Case of Australian Gold'. *The Journal Of Economic History* 44 (1984): 1049.

of exports, analysing how rapid economic growth spurred by gold mining restructured the macroeconomy. Second, I discuss gold's indirect role in reshaping the socio-political structure of Australia to foster more egalitarian economic outcomes. This particular phenomenon was largely attributable to the structural differences between Australia's gold mining and pastoral sectors.

Contributions of the gold industry to macroeconomic restructuring in the nineteenth century

The combination of gold becoming Australia's largest export and the sheer rapidity of the GDP, wage and population growth it created delivered wealth that significantly changed the trajectory and stability of the Australian economy.[3] To understand the mechanisms by which these changes occurred, the context of gold's rise to prominence must first be understood. I then explore the two primary ways in which gold's rise strengthened Australia's economy: increasing economic activity that sustained the growth secondary industries; and securing stable economic growth.

Context of the gold rush

Gold occupied the driver's seat in the Australian economy in the latter half of the nineteenth century. To usurp wool as the continent's dominant industry, gold relied upon being a far more lucrative industry for working-class labourers than the competing wool industry. In the early days of the rush, the gold industry's attractiveness was attributable to the industry's low barriers to entry, and high returns on labour.[4] Alluvial diggers needed almost no capital or land resources, and nominal licensing fees permitted easy industry entrance.[5] Spectacularly, the alluvial digger of 1851 could expect to see their annual income quintuple within 12 months.[6]

Wage increases in the gold industry spread to other industries because employers were faced with the economic reality that unless they increased wages, their employees would leave, as they were better off trying their luck on the goldfields. Improved wages consolidated to around a 70 per cent increase in real wages by the end of the decade.[7] This is contrasted against the wool industry, where the climatic constraints of the pastoral frontier, reliance on convict labour and the squatter oligopoly created by the

[3] La Croix, 'Property Rights and Institutional Change', 9.

[4] Maddock and McLean, 'Supply-Side Shocks', 1052.

[5] Ibid, 1054.

[6] Ibid, 1053.

[7] Ibid.

1847 Lease Agreements limited growth and wages.[8] The allure of gold led labour from around the colonies and overseas to flood the goldfields.[9]

Spurred by recurring discoveries of gold deposits, capital began to flow into the industry, as companies became involved in sophisticated mining methods like deep alluvial gold mining and gold-bearing quartz extraction.[10] The inelasticity of gold prices, due to the international economy's dependence on the metal for the gold standard of currency, and the industry's independence from climatic conditions such as drought, fostered confidence in labour and investors alike. These factors sustained the heavy inflow of labour and capital into the mining industry, allowing the industry to crystallise into the engine room of Australia's most socially influential boom.

Increased economic activity stimulating secondary industry growth

The gold rush in Australia stimulated two types of secondary industries in Australia. First, it supported industries that complimented the expanding population and increased wealth of the working class. Second, it supported industries that directly backed the mining sector. The gold industry's deliverance of increases to average wages, a population boom creating greater demand for manufactured goods, and the increase in Australia's import capacity as its exports exponentially rose led to increased aggregate demand for goods.[11] The Australian working class experienced the best living conditions of any working class in the world during this era,[12] with wages rising across all sectors due to labour shortages in many non-substitutable industries such as construction. The effect of this was that the per capita consumption of imports doubled from 1851 to 1860, causing short-term shocks in industries like manufacturing but leading to long-term growth and stability. To highlight this effect, rapid wage growth in the 1850s created immediate decline in import-competing manufacturing industries such as beer and clothing.[13] This was because international firms paid lower wages and were therefore able to produce cheaper goods, granting them a competitive advantage over domestic producers.[14] However, this decline was offset by the 'spending effect' of wage increases and the population boom generating increased demand for non-tradable domestic manufacturing like construction.[15] Construction as a subset of manufacturing is not

[8] Paul Cashin and C John McDermott, '"Riding on the Sheep's Back": Examining Australia's Dependence on Wool Exports'. *Economic Record* 78 (2002): 249–63.

[9] Maddock and McLean, 'Supply-Side Shocks', 1054

[10] La Croix, 'Property Rights And Institutional Change', 12.

[11] Maddock and McLean, 'Supply-Side Shocks', 1054.

[12] La Croix, 'Property Rights And Institutional Change', 13.

[13] Ibid.

[14] Ibid, 9.

[15] Ibid, 20.

subject to competition from imports and experienced a net improvement due to this demand spike. Nonetheless, the rise of construction required a gross reallocation of resources, which took time.

While the early years of the gold rush actually harmed manufacturing in Australia, once the industry evolved manufacturing grew as an employer and exporter. Considering the counterfactual in which gold never boomed, demand for manufactured goods would most likely have remained low without real wage increases. Some might counter-argue that the role of government in stimulating manufacturing by creating tariff walls was never contingent on a gold rush, and it could be said that manufacturing may have grown as an export industry regardless. However, gold provided labour to facilitate a manufacting industry, increased demand by growing the consuming class, and provided the political cover to introduce tariffs which had long been opposed by wool traders.[16] Therefore, gold can be largely credited for manufacturing growth in the Australian economy, particularly in regard to construction.

Gold's immense growth also birthed industries that serviced the gold industry and secured Australia's international economic position. The industries that grew with the gold rush include banking, transportation across both land and sea, and the manufacturing of equipment used for deep mining and the processing of gold contained in quartz, when the alluvial era ended.[17] By contrast, the previous hegemon of wool was not capital intensive, and therefore there was little demand for either manufacturing farming equipment or for finance as provided for by the banking sector.[18] Of course, later technological advancements to incorporate more capital equipment into the agricultural sector, and mining of other minerals would have also fostered the growth of these supporting industries. In the nineteenth century, however, gold was the most demanded international good, meaning that only gold could deliver growth on the scale observed in nineteenth century Australia. Further, in creating increased aggregate demand that positively translated into an expansion of secondary industries that reflected Australia's increased wealth, there was potential for growth in the gold industry as the alluvial era ended.

Securing sustainable economic growth

Gold contributed to Australian economic stability by increasing the size of the labour force, and by providing Australia with an export that was practically immune to market shocks. Gold gave Australia an economy that, for the first time, was of substantial enough size to avoid being derailed by exogenous shocks in international markets. Following the disastrous 1840s recession that was almost entirely attributable to the surplus in the British wool market,[19] Australian colonial governments became dedicated to the goal of increasing its population, with over 270,000 assisted immigrants, and 550,000 immigrants

[16] Cashin and McDermott, '"Riding on the Sheep's Back"', 257.

[17] Maddock and McLean, 'Supply-Side Shocks', 1054.

[18] Cashin and McDermott, '"Riding on the Sheep's Back"', 255.

[19] Ibid, 253.

overall arriving on Australian shores between 1850 and 1860.[20] This helped in developing a more robust economy, highlighted by the response of the economy to the mild decline of gold exports in the 1870s when wool boomed.[21] Here, the population boom and transition to smaller farming plots that followed the gold rush provided the labour base for the economy to quickly switch focus, and subsequently Australian sheep populations and wool exports both more than doubled in this decade.[22]

In the nineteenth century there was virtually 'unquenchable demand' for the strategic commodity of gold in international markets that protected it from shocks.[23] This set gold further apart from the preceding dominant commodity: wool.[24] Not only was gold unthreatened by climatic changes like drought, but, as noted above, demand for it was virtually ineleastic. As the nineteenth century came to a close, the independence of gold's value from all other economic factors was shown, as the 1890s gold rush in Western Australia staved off a global and domestic depression. As the gold-driven economy returned the current account to surplus by the early 1900s, it increased Western Australia's population fourfold, and doubled mining's contribution to Australia's GDP, despite weaker international markets.[25] Contrasted against an Australian economy where gold hypothetically did not exist, the capacity of the pastoral frontier was constrained by the need for rainfall to be above 10 mm annually,[26] meaning that there was little expansion possible after the 1870s demand and production spike.[27]

If gold were not discovered, Australia's economy would still be at the mercy of the volatile agriculture market. It would further be condemned to remaining labour scarce, thus preventing an expansion in secondary industries. It cannot be said that other minerals discovered in Australia like copper or coal carried gold's capacity to act as a beacon for labour and capital, as they, like wool, were volatile markets not tied to any currency. Therefore, gold must be almost solely credited for its role in stabilising, growing and diversifying Australia's economy in the latter half of the nineteenth century. Gold's influence was far-reaching and spread beyond the economic sphere, with the industry shaping the foundations of a more egalitarian society as Australia's population soared and the twentieth century dawned.

---

[20] Ian McLean, *Why Australia Prospered* (Princeton, Princeton University Press: 2012), ch. 5.

[21] Ibid.

[22] Ibid.

[23] Cashin and McDermott, '"Riding on the Sheep's Back"', 253.

[24] Ibid.

[25] Irving Kravis, 'Trade as a Handmaiden of Growth: Similarities Between the Nineteenth and Twentieth Centuries'. *The Economic Journal* 80 (1970): 857.

[26] Ian McLean, *Why Australia Prospered*, ch. 5.

[27] Ibid.

Long-term socio-political effects of gold on Australia's development as a nation-state

Gold's rise to prevalence in the Australian economy also created a socio-political culture that cultivated egalitarian outcomes. Adopting a normative lens, the cultural benefits can be specifically categorised as breaking the squatting oligarchy, avoiding the resource curse, distributing wealth equally among workers, and politically empowering the working class.

Breaking the squatting oligarchy

After crossing of the Blue Mountains in 1819, emancipists and free settlers flocked to the inland plains to capitalise on the immense opportunity in the wool industry.[28] The wool industry was one of the only colonial exports to truly flourish in face of capital and labour scarcity and became Australia's largest export by 1830.[29] This was because the capital (sheep) reproduced and the farming was land (rather than labour) intensive. Despite the government's dislike for private citizens exploiting Crown land without providing consideration for their use of the land ('squatting'), the authority's hands were largely tied by the immense political and economic clout of the squatting class. The recession of 1840 revealed the significance of wool to the Australian economy, and the failing of the 1844 lease agreements to distribute land rights more fairly reflect this.[30] The consequence of this was that, by 1860, 70 million hectares of land was shared between only 2,000 squatters.[31]

This culture, where squatters held the balance of power, fostered negative social conditions. The squatters' strict property rights to the land and capital components of the economy's leading staple produced an oligarchy. The high barriers to entry into the industry created severe income inequality, which not only is an unfair outcome, but reduces spending and access to imports and foreign capital that could be used to start new business ventures. Furthermore, the squatters used their political clout to prolong the convict era, due to their reliance on unfree shepherding labour.[32] Squatters were against tariffs that could foster manufacturing but would harm wool exports, and opposed suffrage, as they held strong influence in the bureaucratic Legislative Council of NSW.[33]

The discovery of gold in Victoria particularly was responsible for breaking the squatter oligarchy. This was not because gold drew labour from the paddocks to the gold fields, as the advent of fencing largely

---

[28] Cashin and McDermott, '"Riding on the Sheep's Back"', 259.

[29] Ibid, 254.

[30] Morris Altman, 'Staple Theory and Export-Led Growth: Constructing Differential Growth'. *Australian Economic History Review* 43 (2003), 245–50.

[31] Cashin and McDermott, '"Riding on the Sheep's Back"', 262.

[32] Ibid.

[33] Maddock and McLean, 'Supply-Side Shocks', 1058.

circumvented the need for shepherds.[34] Rather, gold incited an immigration boom that tripled Australia's population between 1851 and 1860: this was the factor responsible for breaking the squatting class's political clout.[35] For the first time in Australian history, there was the labour required for land-intensive farming and a pressing need to accommodate the rapidly growing population. Thus, the Settlement Acts of 1861 were enacted, creating smaller farming plots, promoting broader ownership and more land-intensive farming practices.[36] With their land and economic hegemony finally stripped, the era where pastoralists wielded enormous political clout was at end, paving a way for a more progressive age where Australia achieved social milestones like broad male suffrage and minimum wages.

Avoiding the resource curse: A comparison with Argentina

History teaches us that young countries with plentiful deposits of natural resources often fall victim to the resource curse: rapid wealth growth in small pockets of the economy that facilitates oligarchical political outcomes.[37] Gold helped Australia avoid the resource curse; this can be compared with Argentina, which, like Australia, was an economy with high resource stocks but labour scarcity in the nineteenth century. Like Australia, Argentina relied heavily on its rural sector, and its export staples of wheat and beef.[38] Applying staple theory,[39] the strict land rights afforded to agriculturalists in Argentina led to a gross concentration of wealth that led to the same oligarchical political culture that a politically powerful squatting class threatened to deliver to colonial Australia. In Argentina, this concentration of political power culminated in the National Autonomist Party, also known as PAN, emerging as the dominant Argentinian political party. PAN effectively suppressed other political movements in order to preserve the immense wealth concentration and broad economic growth provided by canvassing free trade policy and high barriers to entry in these staple industries.[40] These policies bear a strong resemblance to the canvassing free trade policies implemented by NSW's colonial parliament at the bidding of the economically powerful squatters. However, the overriding influence of gold mining in Australia helped end the rising of a squatting oligarchy, while Argentina's political monopoly was eventually broken in the Barings crisis of the 1890s by the rise of a Radical Party. This Radical Party was met with great resistance by soon-to-be dethroned pastoralists and is broadly agreed to have set back the Argentinian economy for many decades.[41] By contrast, the rise of gold in Australia prevented this reality from ever

---

[34] Maddock and McLean, 'Supply-Side Shocks'.

[35] Ibid, 1049.

[36] Ibid.

[37] Altman, 'Staple Theory and Export-Led Growth'.

[38] Paula Alonso, *Between Revolution and the Ballot Box* (Caimbridge, UK: 2000), 3.

[39] Altman, 'Staple Theory and Export-Led Growth', 245–50.

[40] Alonso, *Between Revolution and the Ballot Box,* 5.

[41] Ibid, 9.

eventuating, with the political clout of the squatting class diminishing as the economic might of the mining industry dominated the economy. It is worth noting that, while Argentina became independent in 1853,[42] the British Government still held sovereignty in 1860s Australia. While both colonial and Argentinean policymakers stood to benefit from the continued boom of the pastoral sector, historians have suggested that the British Government may have intervened to prevent squatters from exerting too much political power in Australia.[43] Notwithstanding this, gold certainly played a role in breaking a predictably dangerous squatting oligarchy, through facilitating the Settlement Acts, and delivering a broad male suffrage following the Eureka Stockade in 1854.

Fair distribution of economic opportunity

Another way in which gold contributed to more egalitarian social outcomes in the Australian economy was the industry's low barriers to entry, making it more accessible to everyone than wool. Contrasted against wool, which rose to prominence in a period of colonial history where free men who were allowed to export were outnumbered by convicts, gold arose in an era where unfree labour populations were dwindling. Furthermore, while land in the 1850s was becoming scarcer as the squatting class consumed it for wool exports, licenses for alluvial gold mining were uncapped.[44] This led to labourers from a range of sectors converging on the goldfields, a sector which, even after equilibration in the late 1850s, provided higher real wages to diggers than what they had previously received.[45]

A secondary consequence of the gold rush was that the economy transitioned to a more wage-based economy. This happened with the rise of manufacturing, and also non-alluvial gold mining where companies began to enter the industry in the 1860s.[46] These opportunities provided greater opportunity for working-class Australians to climb social ladders and experience income security, contrasted against the previously wool-driven economy.

Gold's contribution to the labour movement in Australia

Gold's role in transitioning the economy to a more wage-based economy can also be identified as an origin of the labour movement in Australia, and the emphasis placed on protecting workers' rights by the Australian political left. As discussed above, an immediate effect of the gold rush was labour shortages across other sectors such as construction and pastoralism.[47] The consequence of this was an increase in the price of labour (wages), which subsequently required early gold mining companies such as BHP to

[42] Ibid, 7.

[43] Maddock and McLean, 'Supply-Side Shocks', 1064.

[44] Altman, 'Staple Theory And Export-Led Growth', 237.

[45] Maddock and McLean, 'Supply-Side Shocks', 1054.

[46] Ibid, 1065.

[47] Ibid, 1053.

offer wages that were competitive with other sectors.[48] Suddenly, the Australian working class was experiencing the best working conditions of any working class in the world.[49] By the 1890s, this experience consolidated into a collective sense of entitlement to such working rights. Strikes, campaigns for eight-hour working days and other worker rights coloured the depression of the 1890s in Victoria, leading to the original labour movement combining with the Protectionist Party to form the predecessor to the modern Australian Labor Party (ALP).[50]

The development of the ALP is particularly attributable to the gold rush when considering the Eureka Stockade, which, by birthing democracy in Australia, and 'symbolically fighting for the rights of every man'[51] gave scope for political parties to appeal to the working class. This is different from the previous colonial bureaucracy that clearly favoured individuals and groups with business clout. The crux of Australia's political left labour movement origins is not necessarily reflected in any economic statistics.[52] Where this history becomes relevant, however, is in attributing Australia's rich history of industrial relations and unionism to the nature of its political left, which is distinct from a party that was forged in the politicisation of various social and political ideals, such as the United States' Democratic Party.[53] It is plausible that, if not for the intervention of the gold rush, the Australian political left may have arisen in the form of campaigns for radical social restructuring, as happened in Argentina – unlike the modern ALP, which was formed through campaigning for the preservation and continuation of the broadly high working conditions provided by a gold-influenced economy in Australia.[54]

Conclusion

To conclude, the effect of gold mining in altering the Australian economy's development in the nineteenth century was substantial. It brought growth that was so grossly accelerated that it increased Australia's wealth to a level that was world-leading, and unforeseeable without gold's influence. This growth was so rapid that it permeated all facets of the economy, especially manufacturing and pastoralism, resulting in massive restructures that affected results and human experiences in these sectors. Finally, gold's role in forging Australia's political and social identity leaves a legacy of fierce protection of the working class, which is still effectual today. While gold's impact on the Australian export economy is no longer substantial, the social and political legacy of the once-hegemonic gold industry can be remembered as the

[48] Maddock and McLean, 'Supply-Side Shocks', 1055.

[49] Ibid.

[50] Kravis, 'Trade as a Handmaiden of Growth', 862.

[51] *The Age,* May 8, 1855.

[52] Ibid.

[53] Alonso, *Between Revolution and The Ballot Box*, 7.

[54] Kravis, 'Trade as a Handmaiden of Growth', 857.

bridge between Australia's booming but fragile colonial economy and the robust, diverse and labour-supportive economy of modern Australia.

Acknowledgements

I would like to thank Professor Frank Bongiorno for his dedicated teaching, enchanting wit and for instilling in me an appreciation of the lessons that history has to teach. I also extend my sincere gratitude to Betty Fuller and Christine Te Lindert for fostering my love of writing.

Bibliography

Alonso, Paula. *Between Revolution and the Ballot Box*. 1st ed. Cambridge, UK: Cambridge University Press (2000): 1–14. doi.org/10.1017/CBO9780511572685.

Altman, Morris. 'Staple Theory and Export-Led Growth: Constructing Differential Growth'. *Australian Economic History Review* 43 (2003): 230–55. doi.org/10.1046/j.1467-8446.2003.00053.x.

Cashin, Paul, and C John McDermott. '"Riding on the Sheep's Back": Examining Australia's Dependence on Wool Exports'. *Economic Record* 78 (2002): 249–63. doi.org/10.1111/1475-4932.00055.

Kravis, Irving B. 'Trade as a Handmaiden of Growth: Similarities Between the Nineteenth and Twentieth Centuries'. *The Economic Journal* 80 (1970): 850–72. doi.org/10.2307/2229902.

La Croix, Sumner J. 'Property Rights and Institutional Change During Australia's Gold Rush'. *Explorations In Economic History* 29 (1992): 204–27. doi.org/10.1016/0014-4983(92)90011-K.

Maddock, Rodney, and Ian McLean. 'Supply-Side Shocks: The Case of Australian Gold'. *The Journal Of Economic History* 44 (1984): 1047–67. doi.org/10.1017/S0022050700033088.

McLean, Ian. *Why Australia Prospered.* 1st ed. Princeton: Princeton University Press (2012): ch 5.

The Age, May 8, 1855. 'The Eureka Stockade'. p. 3: trove.nla.gov.au/newspaper/article/154859784?searchTerm=Eureka%20Stockade&searchLimits=dateFrom=1853-01-01|||dateTo=1861-01-01.

The autochthonous development and evolving approach to unjust enrichment by the High Court in Australia

JOSHUA LING

Abstract

Although the High Court had once been lauded by English unjust enrichment scholars for its percipience in its early willingness to recognise the concept of unjust enrichment in 1987, its development and treatment of unjust enrichment has since been the subject of academic and legal controversy. This paper attempts to navigate and interrogate those controversies. It traces, examines and evaluates the evolving approach to unjust enrichment by the High Court from 1987 until today. In so doing, it proffers two observations. First, the High Court has developed and emphasised a conscience-based approach towards restitution, which has not been without contention. Second, despite the common refrain that the Australian and English courts have departed in almost incompatible ways with respect to their jurisprudential understanding and judicial methodology towards unjust enrichment, there is ostensibly greater substantive consistency between the two jurisdictions today than has been hitherto claimed.

Introduction

1987 and 1991 marked watershed moments when unjust enrichment was recognised in Australia and England.[1] Since then, however, it has been claimed that the two jurisdictions have diverged in irreconcilable ways.[2] More remarkably, Australia's approach to unjust enrichment has been censured as being 'in a sorry state'.[3] Given such criticisms, an evaluation of the High Court's approach is both apposite and timely.

[1] *Pavey & Matthews Pty Ltd v Paul* (1987) 162 CLR 221 ('*Pavey*'); *Lipkin Gorman v Karpnale Ltd* [1991] 2 AC 548 ('*Lipkin Gorman*').

[2] See, eg, Paul Finn, 'Common Law Divergences' (2013) 37 *Melbourne University Law Review* 509, 512, 520.

[3] Andrew Burrows, *The Law of Restitution* (Oxford University Press, 3rd ed, 2011) 43.

This paper begins with a historical background to restitution law and unjust enrichment. Thereafter, it traces the autochthonous development of unjust enrichment by the High Court, and makes two observations. First, the High Court has adopted a conscience-based methodology which was driven by concerns that unjust enrichment was illusory.[4] It is, however, questionable whether a conscience-based approach to the common money counts is truly a more cogent basis than unjust enrichment. Second, the disagreements underlying the unjust enrichment debate revolve largely around questions of taxonomy and form. Critics of the High Court's conscience-based approach tend to fixate on its outward inconsistencies with England and other common law jurisdictions. Yet, beyond the dissimilarities in judicial methodology, there is arguably little significant difference in substantive outcomes between the Australian and English approaches. Both approaches may, for instance, be explained by unjust enrichment rationale.[5] Ultimately, this essay suggests that when evaluated as a matter of substance, there is today greater similarity between Australia's and England's approaches to unjust enrichment than has been previously acknowledged.

A brief history of the law surrounding restitutionary claims

Restitution law

Since time immemorial, Roman law has provided restitutionary remedies that reverse a defendant's gains as opposed to providing compensation for losses arising from some wrongdoing.[6] Similarly, restitutionary claims could historically be brought in England via several forms of action at common law. These included the writs of debt and account,[7] which hid the restitutionary nature of the claim beneath a bare plea that money was owed as a debt or must be accounted for.[8] The forms of actions were extended following *Slade's case*,[9] whereby it became permissible for restitutionary claims to be pleaded on assumpsit – the action of

[4] See Paul Finn, 'Equitable Doctrine and Discretion in Remedies' in William Cornish, Richard Nolan, Janet O'Sullivan and Graham Virgo (eds), *Restitution, Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, 1998) 251, 252.

[5] Michael Bryan, 'Peter Birks and Unjust Enrichment in Australia' (2004) 28 *Melbourne University Law Review* 724, 726.

[6] Kit Barker and Ross Grantham, *Unjust Enrichment* (LexisNexis Butterworths, 2nd ed, 2018) 2; James Edelman and Elise Bant, *Unjust Enrichment in Australia* (Oxford University Press, 2006) 79.

[7] Barker and Grantham, above n 6.

[8] James Edelman, 'Australian Challenges for the Law of Unjust Enrichment' (Speech delivered at the Summer School, University of Western Australia, 24 February 2012) 3.

[9] (1648) Style 138.

which was premised upon the breach of a fictional implied contract.[10] Unfortunately, the formulaic nature of these claims meant there was no impetus for the early common law courts to articulate a more comprehensive theory for the various common money counts.[11] This state of affairs came to a head in *Moses v Macferlan*.[12] In that case, Lord Mansfield faced a doctrinal deficiency underpinning the action for money had and received, and thus acknowledging that the nature of the case was not one where such actions had traditionally laid, his Lordship supplemented the fiction of implied contract with principles of 'natural justice and equity'.[13] This, as shown below, has had lasting effects for unjust enrichment in Australia; particularly in the High Court's development of a conscience-based approach to unjust enrichment claims.

Unjust enrichment

Moving into the twentieth century, the fiction of implied contract became increasingly curtailed,[14] and would ultimately be discarded.[15] Instead, unjust enrichment was first raised as a possible explanation for money had and received in the House of Lords in 1943,[16] and by 1954, had even become a cause of action in Canada.[17] In the latter half of the century, academic interest quickly burgeoned over the possibility of ridding the remnants of the old forms of action for restitution, at law and equity, and of unifying them under a singular theory and framework of unjust enrichment.[18]

Today, it is commonly accepted that a successful claim in unjust enrichment requires a four-part inquiry. First, the defendant must have benefited or have been enriched; second, the enrichment must be at the expense of the plaintiff; third, the enrichment must be unjust; and fourth, there must be no defences that are available, such as a change of position.[19] Such a framework is premised upon an 'event-based', as opposed

---

[10] J Baker, 'The History of Quasi-Contract in English Law' in William Cornish, Richard Nolan, Janet O'Sullivan and Graham Virgo (eds), *Restitution, Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, 1998) 37, 39, 41–2, 53–5.

[11] Ibid 39.

[12] (1760) 2 Burr 1005.

[13] Ibid 1012; William Gummow, 'Moses v. Macferlan 250 Years On' 68(3) *Washington and Lee Law Review* 881, 883–4.

[14] See *Sinclair v Brougham* [1914] AC 398, 452 (Lord Sumner).

[15] *Pavey* (1987) 162 CLR 221; *Westdeutsche Landesbank Girozentrale v Islington LBC* [1996] AC 669, 710 ('*Westdeutsche*').

[16] *Fibrosa Spolka Akcyjna v Fairbairn Lawson Combe Barbour Ltd* [1943] AC 32, 61 (Lord Wright).

[17] *Deglman v Guaranty Trust of Canada* [1954] SCR 725.

[18] See Justice Keith Mason, 'Where has Australian Restitution Law Got To and Where is it Going?' (2003) 77 *Australian Law Journal* 358, 359; Robert Goff and Gareth Jones, *The Law of Restitution* (Sweet & Maxwell, 1st ed, 1966) 5.

[19] See Charles Mitchell, Paul Mitchell and Stephen Watterson (eds), *Goff & Jones The Law of Unjust Enrichment* (2016, 9th ed, Sweet & Maxwell) 8; *Banque Financiere de la Cite v Parc (Battersea) Ltd* [1999] 1 AC 221, 227 (Lord Steyn); *Benedetti v Sawiris* [2013] UKSC 50 (17 July 2013) [10]; *Crown Prosecution Services v Eastenders Group* [2014] UKSC 26 (8 May 2014) [102]; *Menelaou v Bank of Cyprus* [2015] UKSC 66

to a 'response-based', taxonomical understanding of private obligations.[20] For example, Birks originally argued that unjust enrichment was one category, out of four, that covered every legal event which triggered a restitutionary remedy.[21] This 'generic conception' would be further subdivided into 'restitution from wrongs', and 'autonomous unjust enrichment' not arising from wrongs.[22] However, this broad conceptualisation of unjust enrichment – as being coterminous with restitution law – was heavily criticised as overly encompassing and therefore internally incoherent.[23] Consequently, Birks formulated a narrower theory which limited unjust enrichment to 'all events materially identical to the mistaken payment of a non-existent debt'.[24]

On Birks' subsequent and narrower view, unjust enrichment becomes a 'subset' of restitution law, and is decoupled from the remedy of restitution.[25] This view was better received, as restitution had since been recognised to arise from legal events *outside* of unjust enrichment too – such as from wrongs or property rights.[26] Ostensibly, the English Supreme Court recently adopted and affirmed the narrower view when it branded the concept of unjust enrichment with cases involving the correction of 'normatively defective transfers of value'.[27] Arguably, the High Court took the narrow view too when it held that the 'unjust factors' within 'unjust enrichment' are concerned with vitiated intention, and not wrongdoing.[28] Nevertheless, in addition to understanding the background history between restitution law and unjust enrichment, it is perhaps more critical to consider more carefully the High Court's particular evolving approach to unjust enrichment in Australia.

(4 November 2015) [18]; *Commissioners for her Majesty's Revenue and Customs v The Investment Trust Companies (in liq)* [2017] UKSC 29 (11 April 2017) [24] ('*Revenue and Customs*').

[20] See Peter Birks (ed), *The Classification of Obligations* (Clarendon Press, 1997) 17–20.

[21] Peter Birks, *An Introduction to the Law of Restitution* (Clarendon Press, 1985) 16–17; see also Goff and Jones, above n 18, 5.

[22] Robert Stevens, 'Is there a Law of Unjust Enrichment?' in Simone Degeling and James Edelman (eds), *Unjust Enrichment in Commercial Law* (Thomas Reuters, 2008) 11, 14.

[23] Ibid 14–16; see especially Steve Hedley, 'Unjust Enrichment' (1995) 54(3) *Cambridge Law Journal* 578.

[24] Peter Birks, *Unjust Enrichment* (Clarendon Press, 2nd ed, 2005) 1.

[25] Ibid 17.

[26] Ibid; James Edelman and Elise Bant, *Unjust Enrichment* (Hart Publishing, 2nd ed, 2016) 19; Graham Virgo, 'What is the Law of Restitution About?' in William Cornish, Richard Nolan, Janet O'Sullivan and Graham Virgo (eds), *Restitution, Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, 1998) 305, 307; see also *Sempra Metals Ltd v Her Majesty's Commissioners for Inland Revenue* [2007] UKHL 34 (18 July 2007) [116] (Lord Nicholls).

[27] *Revenue and Customs* [2017] UKSC 29 (11 April 2017) [42]–[43]; see also *Lowick Rose LLP (in liq) v Swynson* [2017] UKSC 32 (11 April 2017) [22]; cf *Lipkin Gorman* [1991] 2 AC 548, 578 (Lord Goff).

[28] *Farah Constructions Pty Ltd v Say-Dee Pty Ltd* (2007) 230 CLR 89, 156 [150] ('*Farah Constructions*'); see also James Edelman, 'A Principled Approach to Unauthorised Receipt of Trust Property' (2006) 122 *Law Quarterly Review* 174, 177–8.

Autochthonous development in Australia

Broadly speaking, Australia's treatment of unjust enrichment might be divided into three epochs. The first epoch began in 1987 when, without precedent, the High Court dispensed with the fiction of implied contract for restitutionary claims, and explicitly recognised unjust enrichment as a 'unifying legal concept' in *Pavey*.[29] In that case, contract could not form a 'juristic explanation' of the respondent's enrichment because of her repudiatory breach.[30] Thus, the appellant builders were entitled to restitution on the basis of *quantum meruit*, which in turn, was said to be 'based on unjust enrichment'.[31] Immediately thereafter, several lower court decisions attempted to follow suit by embracing unjust enrichment.[32] This was despite a gaping uncertainty as to whether unjust enrichment was a legal principle in its own right, or something else.[33] A high-water mark was reached in 1993 when Deane and Dawson JJ suggested there could even be an 'action in unjust enrichment'.[34] However, although this proposition was soon decisively rejected by the High Court,[35] greater relevance was made for unjust enrichment with the removal of the restitutionary bar for mistakes of law, the creation of a structured approach to unjust enrichment based upon unjust factors (such as mistake, duress or illegality),[36] and the incipient recognition of the change of position defence.[37]

A second epoch arrived at the turn of this century, which saw the High Court retreat from the concept of unjust enrichment, and which involved the development of a conscience-based approach towards restitutionary claims.[38] In *Roxborough*,[39] Gummow J cast much doubt on the continued relevance of unjust

---

[29] (1987) 162 CLR 221, 227–8, 256.

[30] Bryan, 'Peter Birks and Unjust Enrichment in Australia', above n 5, 728.

[31] *Pavey* (1987) 162 CLR 221, 227–8; cf Romauld Andrew, 'The Fabrication of Unjust Enrichment in Australian Law: Pavey & Matthews v Paul Reassessed' (2010) 26 *Building and Construction Law Journal* 314.

[32] See, eg, *National Mutual Life Association of Australasia Ltd v Walsh* [1987] 8 NSWLR 585, 595; *Public Trustee v Fraser* [1987] 9 NSWLR 433, 443; *Nepean District Tennis Association Inc v Penrith City Council* (1988) 66 LGRA 440, 448–9.

[33] Warren Swain, 'Unjust Enrichment and the Role of Legal History in England and Australia' (2013) 36(3) *University of New South Wales Law Journal* 1030, 1041.

[34] *Baltic Shipping Co v Dillon* (1993) 176 CLR 344, 379 ('*Baltic Shipping*').

[35] *David Securities Pty Ltd v Commonwealth Bank of Australia* (1992) 175 CLR 353, 406 ('*David Securities*'); *Hill v Van Erp* (1997) 188 CLR 159, 239.

[36] See Michael Bryan, 'Mistaken Payments and the Law of Unjust Enrichment: David Securities Pty Ltd v Commonwealth Bank of Australia' (1993) 15 *Sydney Law Review* 461, 474–5; see also Chief Justice Allsop, 'Restitution: Some Historical Remarks' (2016) 90 *Australian Law Journal* 561, 577.

[37] *David Securities* (1992) 175 CLR 353; see also Stevens, above n 22, 13.

[38] See *Roxborough v Rothmans of Pall Mall Australia Ltd* (2001) 208 CLR 516 ('*Roxborough*').

[39] Ibid.

enrichment in Australia. His Honour held that while the fiction of implied contract was conclusively rejected by *Pavey*,[40] that case did not identify 'a satisfactory doctrinal basis' for the common money counts.[41] Instead, though unjust enrichment might be useful in furthering 'legal inquiry' as a 'unifying legal concept', it could 'contrive legal analysis' if directly applied as a legal principle since it was apt to conceal the particular responsibilities and relationships which provide the legal and policy basis upon which restitution is granted.[42] Unjust enrichment was an 'all-embracing theory of restitutionary rights'.[43] Its acceptance would lead to 'top-down reasoning' upending common law judicial methodology, which 'may distort well settled [equitable] principles' and remedies.[44] Moreover, actions for money had and received were historically laid even against defendants who were never enriched.[45] Thus, his Honour suggested that a conscience-based approach, focusing on the unconscionability in retaining a conferred benefit, was a superior basis for restitution than unjust enrichment.[46] Such criticisms would be repeated by the High Court over subsequent years.[47]

The third epoch, arriving at the start of this decade, has most noticeably been marked by the High Court's softening of its criticisms against unjust enrichment, and return to the structured approach towards unjust enrichment claims as originally expounded upon in *David Securities*.[48] Thus, in a 2012 decision, it was reaffirmed that restitutionary rights were distinct from contractual claims.[49] In so doing, however, the High Court stated that unjust enrichment retained a useful 'taxonomical function referring to [the] categories of cases' which attract restitution.[50] To that extent, unjust enrichment was not the 'all-embracing theory' as

---

[40] (1987) 162 CLR 221.

[41] *Roxborough* (2001) 208 CLR 516, 540 [64].

[42] Ibid 543 [70].

[43] Ibid 544 [72].

[44] Ibid 544–5 [72]–[74]; cf Carmine Conte, 'From Only the 'Bottom-up'? Legitimate Forms of Judicial Reasoning in Private Law' (2015) 35(1) *Oxford Journal of Legal Studies* 1.

[45] *Roxborough* (2001) 208 CLR 516, 543–4 [71], citing *Martin v Pont* [1993] 3 NZLR 25; cf Burrows, *The Law of Restitution*, above n 3, 37.

[46] *Roxborough* (2001) 208 CLR 516, 554 [100].

[47] *Farah Constructions* (2007) 230 CLR 89, 156 [150]–[151]; *Lumbers v W Cook Builders Pty Ltd* (2008) 232 CLR 635, 662 [77]–[78] ('*Lumbers*'); *Bofinger v Kingsway Group Ltd* (2009) 239 CLR 269, 300 [90]–[91] ('*Bofinger*').

[48] (1992) 175 CLR 353, 376; see especially *Equuscorp Pty Ltd v Haxton* (2012) 246 CLR 498, 516 [30] ('*Equuscorp*').

[49] *Equuscorp* (2012) 246 CLR 498, 514–15.

[50] Ibid 516 [30].

previously thought,[51] and may develop the law in 'novel occasions'.[52] Furthermore, although the High Court was divided over whether illegality or the failure of consideration provided the qualifying or vitiating factor in that case, the full bench agreed that the more critical question was whether the allowance of restitution, in circumstances of illegality, would stultify statutory purpose, and hence undermine coherence in the law.[53] Ostensibly, it has been argued that the High Court had, in effect, permitted 'policy-motivated reasons' for restitution that were independent of the existence of a 'qualifying or vitiating unjust factor'.[54] If true, unjust enrichment is not merely a 'sterile exercise in taxonomy', but may have taken on more 'normative force' in Australia than has been explicitly recognised.[55] In the proceeding section of this paper, two observations are made as commentary on the High Court's evolving approach to unjust enrichment.

Observations

Conscience

The first observation that may be made on the High Court's evolving approach to unjust enrichment relates to its apparent reliance upon the concept of conscience. A closer evaluation of the High Court's conscience-based approach is germane given its tendency for controversy between 'equity' and 'restitution' lawyers.[56] Since *Roxborough*,[57] conscience-based reasoning has seen increased prominence in the courts.[58] Most recently, it was reiterated that unjust enrichment (and disenrichment) was 'not the basis for restitutionary relief in' Australia.[59] Instead, restitution is awarded pursuant 'to equitable principles',[60] viz., whether the

---

[51] Ibid; *Australian Financial Services & Leasing Pty Ltd v Hills Industries Ltd* (2014) 253 CLR 560, 595 [74] ('*Hills*'); cf *Roxborough* (2001) 208 CLR 516, 544 [72].

[52] *Equuscorp* (2012) 246 CLR 498, 516 [30].

[53] Ibid 518 [33]–[34] (French CJ, Crennan and Kiefel JJ), 541 [103] (Gummow and Bell JJ), 547–8 [122] (Heydon J).

[54] Elise Bant, 'Illegality and the Revival of Unjust Enrichment in Australia' (2012) 128 *Law Quarterly Review* 341, 344; see also *Woolwich Equitable Building Society v Inland Revenue Commissioners* [1993] AC 70; *Kiriri Cotton Co Ltd v Dewani* [1960] AC 192.

[55] Bant, 'Illegality and the Revival of Unjust Enrichment in Australia', above n 54; Swain, above n 33, 1051; see also Greg Weeks, 'The Public Law of Restitution' (2014) 38 *Melbourne University Law Review* 198.

[56] See, eg, *Westdeutsche* [1996] AC 669, 685 (Lord Goff).

[57] (2001) 208 CLR 516.

[58] See, eg, *Equuscorp* (2012) 246 CLR 498; *Kakavas v Crown Melbourne Limited* (2013) 250 CLR 392; *Pitt v Holt* [2013] 2 AC 108; *Hills* (2014) 253 CLR 560.

[59] *Hills* (2014) 253 CLR 560, 596 [78].

[60] Ibid 597 [78].

retention of money paid to the defendant is 'inequitable'[61] or 'unconscionable'.[62] As aforementioned, the preference for a conscience-based approach stemmed from concerns that unjust enrichment was vague, and so shrouded the nature of liability for restitution.[63] Moreover, a conscience-based approach is consistent with the equitable language of settled case law for money had and received;[64] and the judicial predisposition in Australia for remedies to be guided by 'appropriateness and not *a priori* specification'.[65]

It is nonetheless doubtful that a conscience-based approach is more cogent than the four-part inquiry framework as strictly set out under unjust enrichment theory. First, unconscionability, especially when used as a 'basal principle', may be more ambiguous than unjust enrichment.[66] This is because unlike unjust enrichment, unconscionability has no self-evident content.[67] Conscience's 'guiding criteria'[68] would always require ex post facto enunciation by the courts to be practically understood by lawyers,[69] and it is little consolation that unconscionability 'is not indeterminate' if given time to develop.[70] Indeed, the judicial norms underpinning unconscionability and its function (as a doctrinal rationale or determinant of liability) differ according to the legal doctrine at hand.[71] This may result in 'judicial idiosyncrasy' in its application.[72]

---

[61] Ibid 568 [1], 594 [69].

[62] Ibid 592 [65].

[63] *Roxborough* (2001) 208 CLR 516, 543 [70].

[64] *Moses v Macferlan* (1760) 2 Burr 1005, 1008–12, quoted in *Roxborough* (2001) 208 CLR 516, 548 [83]; see also Ben Kremer, 'Case Comment: Restitution and Unconscientiousness: Another View' (2003) 119 *Law Quarterly Review* 188.

[65] Finn, 'Equitable Doctrine and Discretion in Remedies', above n 4, 266; see *Roxborough* (2001) 208 CLR 516, 545 [75]; see also *Bofinger* (2009) 239 CLR 269, 300–1 [92]–[93]; see, eg, *Trade Practices Act 1974* (Cth) s 87.

[66] J Beatson and G Virgo, 'Case Comment: Contract, Unjust Enrichment and Unconscionability' (2002) 118 *Law Quarterly Review* 352, 354; see also *Garcia v National Australia Bank* (1998) 194 CLR 395, 409 (Gaudron, McHugh, Gummow and Hayne JJ); *ACCC v C G Berbatis Holdings Pty Ltd* (2003) 214 CLR 51, 73 (Gummow and Hayne JJ).

[67] Edelman and Bant, *Unjust Enrichment in Australia*, above n 6, 93 n 130; *Royal Brunei Airlines Sdn Bhd v Tan* [1995] 2 AC 378, 392 (Lord Nicholls).

[68] *Hills* (2014) 253 CLR 560, 576 [16].

[69] Elise Bant, 'Constructive Trusts, Unconscionability and the Necessity for Working Criteria' (2014) 8 *Journal of Equity* 259, 280.

[70] Cf James McConvill and Mirko Bagaric, 'The Yoking of Unconscionability and Unjust Enrichment in Australia' 7(2) *Deakin Law Review* 225, 240.

[71] Pauline Ridge, 'Modern Equity: Revolution or Renewal from Within?' in Sarah Worthington, Andrew Robertson and Graham Virgo (eds), *Revolution and Evolution in Private Law* (Bloomsbury Publishing, 2018) 251, 266.

[72] Robert Boadle, 'Conscience and Unjust Enrichment' (2015) 89 *Australian Law Journal* 641, 650.

Second, a conscience-based approach may also be incompatible with the structured approach as laid down by *David Securities*.[73] Unconscionability's opaqueness lends itself to conceptual confusion.[74] Arguably, this has manifested in the lower courts' attempts at developing an alternative jurisprudence of 'unconscionable retention of benefit'.[75] This may have the effect of threatening unjust enrichment's conceptual coherence by requiring knowledge on the defendant's part, and by subsuming the secondary issue of change of position under the primary question of liability.[76] Additionally, a focus on a defendant's unconscionable *retention* (as opposed to receipt) suggests liability is *contingent* upon the quality of the defendant's conduct.[77] Thus, there is a risk that unconscionability introduces considerations of wrongdoing inconsistent with the strict liability approach taken in *David Securities*.[78] Consequently, because restitution focuses upon the defendant's gain, and not loss caused,[79] the more appropriate remedy for cases like *Roxborough* [80] should arguably be compensation.[81] If understood as a species of wrongdoing, unconscionability may paradoxically need to 'generate new fictions … to support its thesis'.[82] This is because wrongdoings typically require further judicial pronouncement of the antecedent obligation that was breached.[83]

Nevertheless, attempts have been made to strengthen the doctrinal cogency of the conscience-based approach, although they have not been wholly satisfactory. One academic, for instance, argues that unconscionability is an unjust factor grounding restitution for unjust enrichment, and not a wrong unless statutorily prescribed.[84] Others argue that normativity in judge-made law is inevitable, and unconscionability is but an honest way of confronting this reality.[85] More pertinently, the High Court has

---

[73] (1992) 175 CLR 353, 376; see also Bryan, 'Peter Birks and Unjust Enrichment in Australia', above n 5, 733.

[74] See Edelman and Bant, *Unjust Enrichment*, above n 26, 27.

[75] *Heperu Pty Ltd v Belle* (2009) 258 ALR 727; *Ford v Perpetual Trustees Victoria Ltd* (2009) 257 ALR 658.

[76] Bant, 'Illegality and the Revival of Unjust Enrichment in Australia', above n 54, 344; cf Justice Susan Kiefel, 'Lessons from a "Conversation" about Restitution' (2014) 14(2) *Queensland University of Technology Law Review* 1.

[77] Ross Grantham, 'Restitutionary Recovery Ex Aequo et Bono' (2002) *Singapore Journal of Legal Studies* 388, 398–9.

[78] (1992) 175 CLR 353, 376; see also Edelman and Bant, *Unjust Enrichment in Australia*, above n 6, 32–3, 94.

[79] Grantham, above n 77, 400.

[80] (2001) 208 CLR 516.

[81] Grantham, above n 77, 400.

[82] *Roxborough* (2001) 208 CLR 516, 545 [74].

[83] Grantham, above n 77, 401–2.

[84] Michael Bryan, 'Unconscionable Conduct as an Unjust Factor' in Simone Degeling and James Edelman (eds), *Unjust Enrichment in Commercial Law* (Thomson Reuters, 2008) 295–315.

[85] See, eg, Ridge, above n 71, 269–70; Kremer, above n 64, 191–2.

moved to narrow Gummow J's comments in *Roxborough*.[86] Thus, unconscionability is not a 'subjective evaluation of what is fair or unconscionable',[87] but must be understood by reference to 'a qualifying or vitiating factor'.[88] The High Court has even gone so far as to clarify that 'principles of restitution or unjust enrichment can be equated with seminal equitable notions of good conscience';[89] thereby suggesting that unconscionability has no separate function than to describe the presence of unjust factors.[90] Yet, if true, this reduces unconscionability to mere tautology;[91] and, if so, the question remains about the utility of retaining unconscionability as the 'fifth wheel on the unjust enrichment coach'.[92]

Form over substance

A second observation that may be made is that criticisms of Australia's approach to unjust enrichment tend to fixate upon its outward discrepancies with England and other common law jurisdictions.[93] Disagreements tend to be over questions of form – taxonomy, methodology and what unjust enrichment '*should look like*' – than of substance.[94] One illustration is captured by the debates surrounding the role unjust enrichment *ought* to play. For example, Barker writes that unjust enrichment serves four potential roles: it may be a 'classificatory unit'; an 'extrinsic norm'; a legal principle with normative force and legal status; or a cause of action.[95] Similarly, Virgo suggests that unjust enrichment is typically understood in a 'descriptive sense', which simply describes 'a state of affairs where the defendant [obtains] a benefit in circumstances of injustice';[96] or in a 'substantive sense', where unjust enrichment is legally dispositive of a defendant's

[86] (2001) 208 CLR 516.

[87] *Equuscorp* (2012) 246 CLR 498, 518 [32].

[88] Ibid; see also *Hills* (2014) 253 CLR 560, 596 [76].

[89] *Hills* (2014) 253 CLR 560, 576 [16], 595 [74].

[90] Bant, 'Illegality and the Revival of Unjust Enrichment in Australia', above n 54; Edelman and Bant, *Unjust Enrichment in Australia*, above n 6, 94.

[91] Birks, above n 24, 5–6; Edelman and Bant, *Unjust Enrichment*, above n 26, 27.

[92] Burrows, *The Law of Restitution*, above n 3, 37.

[93] See, eg, ibid 35.

[94] Kit Barker, 'Unjust Enrichment: Containing the Beast' (1995) 15(3) *Oxford Journal of Legal Studies* 457, 463; see also Edelman and Bant, *Unjust Enrichment*, above n 26, 15–19.

[95] Kit Barker, 'Understanding the Unjust Enrichment Principle in Private Law: A Study of the Concept and its Reasons' in JW Neyers, M McInnes and S G Pitel (eds), *Understanding Unjust Enrichment* (Hart Publishing, 2004) 79, 84–90.

[96] Virgo, 'What is the Law of Restitution About?', above n 26, 310.

liability to make restitution upon the above-mentioned four-part inquiry.[97] Ostensibly, disagreements with the High Court's approach occur largely at this level of analysis.[98]

Unlike the High Court, unjust enrichment scholars posit that unjust enrichment *ought* to be understood in the substantive sense and given effect via a strict adherence to taxonomy. This is driven by the academic desire for coherence in the law that like cases be treated alike.[99] Such desire is demonstrated by a rigorous taxonomical approach that emphasises the purity of an 'event-based' taxonomy, and direct application of 'unjust enrichment at the claimant's expense [as a] *cause of action*'.[100] Unjust enrichment should be organised as a uniform 'category of claims … whose members respond to the same normative concerns and share the same normative justification'.[101] This is because there is intrinsic good in having a settled taxonomical paradigm which makes the law easier to apply, more accessible, logically transparent and elegant.[102] A taxonomical approach concerns itself with the 'deep structures' of normativity underlying legal doctrine, beyond the 'merely contextual or jurisdictional categories' characterising the old forms of action.[103]

Accordingly, it has been contended that the High Court's apparent preference for traditional pleadings and refusal to apply the four-part inquiry is reminiscent of the old forms of action and should be discouraged. For example, in *Farah Constructions*,[104] the High Court favoured a traditional equitable analysis and fault-based standard for knowing receipt over the unjust enrichment analysis and strict liability approach taken by the Court of Appeal. This conclusion was criticised as undermining legal coherence since it potentially meant having two 'models of restitutionary liability' — in equity and unjust enrichment — applying to 'essentially the same fact pattern'.[105] Similarly, the High Court's emphasis on the need to show work done

---

[97] Ibid 310–11.

[98] Barker and Grantham, above n 6, 12.

[99] Andrew Burrows, 'We Do This at Common Law but That in Equity' (2002) 22(1) *Oxford Journal of Legal Studies* 1, 4.

[100] Burrows, *The Law of Restitution*, above n 3, 4–5, 27 (emphasis added); cf *Bofinger* (2009) 239 CLR 269, 301 [93]; cf *Revenue and Customs* [2017] UKSC 29 (11 April 2017) [41].

[101] Lionel Smith, 'Unjust Enrichment: Big or Small?' in Simone Degeling and James Edelman (eds), *Unjust Enrichment in Commercial Law* (Thomson Reuters, 2008) 35.

[102] Ewan McKendrick, 'Taxonomy: Does it Matter?' in David Johnston and Reinhard Zimmerman (eds), *Unjustified Enrichment: Key Issues in Comparative Perspective* (Cambridge University Press, 1st ed, 2002) 627, 632–8; see also Steve Hedley, 'Rival Taxonomies Within Obligations: Is There a Problem?' in Simone Degeling and James Edelman (eds), *Equity in Commercial Law* (Lawbook Co, 2005) 78, 84–8.

[103] Simone Degeling and Mehera San Roque, 'Unjust Enrichment: A Feminist Critique of Enrichment' 36 *Sydney Law Review* 69, 71.

[104] (2007) 230 CLR 89.

[105] Burrows, *The Law of Restitution*, above n 3, 39.

at the defendant's request for a quantum meruit claim,[106] was denounced as contradictory to unjust enrichment's aims of disaggregating the old forms of action for greater logical transparency.[107]

Yet, focusing on the High Court's outward methodological discrepancies arguably distracts from the anterior question as to why such discrepancies exist,[108] and is apt to ignore the *substantive* similarities between Australia's approach and that taken by other common law jurisdictions. For example, the high-level 'quest for "coherence"' undertaken by unjust enrichment scholars is equally important to the High Court too.[109] Arguably, the High Court's apparently idiosyncratic methodology to unjust enrichment stems from a localised legal history more strongly rooted in equity than England.[110] Australia has a 'strong "preservationist" tradition' for equity owing to the 'late "fusion" of law and equity in New South Wales';[111] and equity continues to be normatively seen as indispensable today.[112] Thus, the taxonomical debates led by unjust enrichment scholars have not influenced the direction of unjust enrichment in Australia as it has in England.[113] Consequently, although coherence would be understood by the High Court and unjust enrichment scholars as a preference for 'doctrinal or conceptual fit to historical fit',[114] this has been expressed in outwardly divergent ways. A clear example is the role of equity within unjust enrichment. To unjust enrichment scholars, coherence means amalgamating restitution at common law and equity into an 'event-based' taxonomy as reflected under a four-part inquiry.[115] To the High Court, however, coherence means assimilating equitable notions of 'good conscience' into the 'fabric of the common law', with equity prevailing.[116]

---

[106] See *Lumbers* (2008) 232 CLR 635.

[107] Burrows, *The Law of Restitution*, above n 3, 41.

[108] Barker, 'Unjust Enrichment: Containing the Beast', above n 94, 463.

[109] Justice Mason, above n 18.

[110] Swain, above n 33, 1049–50.

[111] Barker and Grantham, above n 6, 55; see *Supreme Court Act 1970* (NSW) s 57.

[112] Anthony Mason, 'The Place of Equity and Equitable Remedies in the Contemporary Common Law World' (1994) 110 *Law Quarterly Review* 238, 239.

[113] Swain, above n 33, 1050.

[114] Justice Mason, above n 18; see also *Baltic Shipping* (1993) 176 CLR 344, 376 (Deane and Dawson JJ); *Hill v Van Erp* (1997) 188 CLR 159, 231 (Gummow J); *Tame v New South Wales* (2002) 211 CLR 317, 381 [191] (Gummow and Kirby JJ).

[115] Burrows, *The Law of Restitution*, above n 3, 25–6; Burrows, 'We Do This at Common Law but That in Equity', above n 99, 4–5.

[116] *Roxborough* (2001) 208 CLR 516, 554 [100].

Nevertheless, despite disagreements over judicial methodology, it has been noted that similar substantive outcomes would be reached under either approach.[117] Indeed, the High Court's preference for the old forms of action (e.g. money had and received) can easily be translated into the four-part framework and explained by unjust enrichment rationale.[118] Moreover, the apparent dissimilarities between the High Court's use of unconscionability, and England's preference for the four-part inquiry,[119] disappear when examined as a matter of substance. As above-mentioned, unconscionability has ostensibly been recently clarified as not having an independent role outside an unjust factor.[120] Thus, it may presently be the case, that whereas 'unjust enrichment operates to establish whether the receipt of the enrichment is unconscionable in a principled sense' in England,[121] 'unconscionability can only be interpreted in a principled way' by reference to unjust enrichment in Australia.[122] To distinguish one from the other simply on the basis of methodology is, today, arguably a distinction without difference.[123]

Conclusion

This paper has examined and evaluated the High Court's evolving approach to unjust enrichment since 1987. It began with a brief historical account of restitution law and the concept of unjust enrichment. Thereafter, it traced the autochthonous development of unjust enrichment by the High Court at case law and made two observations. First, the High Court has adopted a conscience-based methodology for cases involving unjust enrichment. This was driven by concerns that unjust enrichment was illusory. It, however, remains open whether a conscience-based approach is a superior basis for restitution than that provided by unjust enrichment. Second, criticisms of the High Court's approach tend to fixate upon its outward inconsistencies with the rest of the common law world. Yet, as this paper suggests, as a matter of substance there ultimately exists, today, greater similarity between Australia's and England's approaches to unjust enrichment than has been previously acknowledged. If this is true, it may now be possible to say, at least,

---

[117] See *Baumgartner v Baumgartner* (1987) 164 CLR 137, 154 (Toohey J); see also Burrows, *The Law of Restitution*, above n 3, 35, 39–40, 42.

[118] See Lionel Wirth, 'Unjust Enrichment: Unifying Concept or Cause of Action?' (2015) 89(5) *Law Institute Journal* 34; see also Edelman and Bant, *Unjust Enrichment in Australia*, above n 6, 85 n 65.

[119] See *Revenue and Customs* [2017] UKSC 29 (11 April 2017) [24], [41].

[120] Bant, 'Illegality and the Revival of Unjust Enrichment in Australia', above n 54; see *Equuscorp* (2012) 246 CLR 498, 518 [32]; *Hills* (2014) 253 CLR 560, 596 [76].

[121] See especially *Revenue and Customs* [2017] UKSC 29 (11 April 2017) [39]–[41].

[122] Graham Virgo, 'Conscience or Unjust Enrichment?: The Emperor's Old Clothes: Australian Financial Services and Leasing Pty Ltd v Hills Industries Ltd' on *Opinions on High* (19 May 2014) <blogs.unimelb.edu.au/opinionsonhigh/2014/05/19/virgo-hills-industries/>.

[123] Ibid; see especially Boadle, above n 72, 653 n 128; Edelman and Bant, *Unjust Enrichment*, above n 26, 13–14.

that the Australian and English approaches to unjust enrichment are no longer so divergent or irreconcilable as popularly believed.

Bibliography

Articles/books/reports

Andrew, Romauld, 'The Fabrication of Unjust Enrichment in Australian Law: Pavey & Matthews v Paul Reassessed' (2010) 26 *Building and Construction Law Journal* 314

Baker, J, 'The History of Quasi-Contract in English Law' in William Cornish, Richard Nolan, O'Sullivan, Janet and Graham Virgo (eds), *Restitution, Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, 1998) 37

Bant, Elise, 'Illegality and the Revival of Unjust Enrichment in Australia' (2012) 128 *Law Quarterly Review* 341

Bant, Elise, 'Constructive Trusts, Unconscionability and the Necessity for Working Criteria' (2014) 8 *Journal of Equity* 259

Barker, Kit, 'Unjust Enrichment: Containing the Beast' (1995) 15(3) *Oxford Journal of Legal Studies* 457

Barker, Kit, 'Understanding the Unjust Enrichment Principle in Private Law: A Study of the Concept and its Reasons' in JW Neyers, M McInnes and SG Pitel (eds), *Understanding Unjust Enrichment* (Hart Publishing, 2004) 79

Barker, Kit and Ross Grantham, *Unjust Enrichment* (LexisNexis Butterworths, 2nd ed, 2018)

Beatson, J and G Virgo, 'Case Comment: Contract, Unjust Enrichment and Unconscionability' (2002) 118 *Law Quarterly Review* 352

Birks, Peter, *An Introduction to the Law of Restitution* (Clarendon Press, 1985)

Birks, Peter (ed), *The Classification of Obligations* (Clarendon Press, 1997)

Birks, Peter, *Unjust Enrichment* (Clarendon Press, 2nd ed, 2005)

Boadle, Robert, 'Conscience and Unjust Enrichment' (2015) 89 *Australian Law Journal* 641

Bryan, Michael, 'Mistaken Payments and the Law of Unjust Enrichment: David Securities Pty Ltd v Commonwealth Bank of Australia' (1993) 15 *Sydney Law Review* 461

Bryan, Michael, 'Peter Birks and Unjust Enrichment in Australia' (2004) 28 *Melbourne University Law Review* 724

Bryan, Michael, 'Unconscionable Conduct as an Unjust Factor' in Simone Degeling and James Edelman (eds), *Unjust Enrichment in Commercial Law* (Thomson Reuters, 2008) 295

Burrows, Andrew, 'We Do This at Common Law but That in Equity' (2002) 22(1) *Oxford Journal of Legal Studies* 1

Burrows, Andrew, *The Law of Restitution* (Oxford University Press, 3rd ed, 2011)

Chief Justice Allsop, 'Restitution: Some Historical Remarks' (2016) 90 *Australian Law Journal* 561

Conte, Carmine, 'From Only the 'Bottom-up'? Legitimate Forms of Judicial Reasoning in Private Law' (2015) 35(1) *Oxford Journal of Legal Studies* 1

Degeling, Simone and Mehera San Roque, 'Unjust Enrichment: A Feminist Critique of Enrichment' 36 *Sydney Law Review* 69

Edelman, James, 'A Principled Approach to Unauthorised Receipt of Trust Property' (2006) 122 *Law Quarterly Review* 174

Edelman, James and Elise Bant, *Unjust Enrichment in Australia* (Oxford University Press, 2006)

Edelman, James and Elise Bant, *Unjust Enrichment* (Hart Publishing, 2nd ed, 2016)

Finn, Paul, 'Equitable Doctrine and Discretion in Remedies' in William Cornish, Richard Nolan, Janet O'Sullivan and Graham Virgo (eds), *Restitution, Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, 1998) 251

Finn, Paul, 'Common Law Divergences' (2013) 37 *Melbourne University Law Review* 509

Goff, Robert and Gareth Jones, *The Law of Restitution* (Sweet & Maxwell, 1st ed, 1966)

Grantham, Ross, 'Restitutionary Recovery Ex Aequo et Bono' (2002) *Singapore Journal of Legal Studies* 388

Gummow, William, 'Moses v. Macferlan 250 Years On' 68(3) *Washington and Lee Law Review* 881

Hedley, Steve, 'Unjust Enrichment' (1995) 54(3) *Cambridge Law Journal* 578

Hedley, Steve, 'Rival Taxonomies Within Obligations: Is There a Problem?' in Simone Degeling and James Edelman (eds), *Equity in Commercial Law* (Lawbook Co, 2005) 78

Justice Kiefel, Susan, 'Lessons from a "Conversation" About Restitution' (2014) 14(2) *Queensland University of Technology Law Review* 1

Justice Mason, Keith, 'Where has Australian Restitution Law Got To and Where is it Going?' (2003) 77 *Australian Law Journal* 358

Kremer, Ben, 'Case Comment: Restitution and Unconscientiousness: Another View' (2003) 119 *Law Quarterly Review* 188

Mason, Anthony, 'The Place of Equity and Equitable Remedies in the Contemporary Common Law World' (1994) 110 *Law Quarterly Review* 238

McConvill, James and Mirko Bagaric, 'The Yoking of Unconscionability and Unjust Enrichment in Australia' 7(2) *Deakin Law Review* 225

McKendrick, Ewan, 'Taxonomy: Does it Matter?' in David Johnston and Reinhard Zimmerman (eds), *Unjustified Enrichment: Key Issues in Comparative Perspective* (Cambridge University Press, 1st ed, 2002) 627. doi.org/10.1017/CBO9780511495519.023

Mitchell, Charles, Paul Mitchell and Stephen Watterson (eds), *Goff & Jones The Law of Unjust Enrichment* (2016, 9th ed, Sweet & Maxwell)

Ridge, Pauline, 'Modern Equity: Revolution or Renewal from Within?' in Sarah Worthington, Andrew Robertson and Graham Virgo (eds), *Revolution and Evolution in Private Law* (Bloomsbury Publishing, 2018) 251

Smith, Lionel, 'Unjust Enrichment: Big or Small?' in Simone Degeling and James Edelman (eds), *Unjust Enrichment in Commercial Law* (Thomson Reuters, 2008) 35

Stevens, Robert, 'Is there a Law of Unjust Enrichment?' in Simone Degeling and James Edelman (eds), *Unjust Enrichment in Commercial Law* (Thomas Reuters, 2008) 11

Swain, Warren, 'Unjust Enrichment and the Role of Legal History in England and Australia' (2013) 36(3) *University of New South Wales Law Journal* 1030

Virgo, Graham, 'What is the Law of Restitution About?' in William Cornish, Richard Nolan, Janet O'Sullivan and Graham Virgo (eds), *Restitution, Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, 1998) 305

Weeks, Greg, 'The Public Law of Restitution' (2014) 38 *Melbourne University Law Review* 198

Wirth, Lionel, 'Unjust Enrichment: Unifying Concept or Cause of Action?' (2015) 89(5) *Law Institute Journal* 34

Cases

ACCC v C G Berbatis Holdings Pty Ltd (2003) 214 CLR 51

Australian Financial Services & Leasing Pty Ltd v Hills Industries Ltd (2014) 253 CLR 560

Baltic Shipping Co v Dillon (1993) 176 CLR 344

Banque Financiere de la Cite v Parc (Battersea) Ltd [1999] 1 AC 221

Baumgartner v Baumgartner (1987) 164 CLR 137

Benedetti v Sawiris [2013] UKSC 50 (17 July 2013)

Bofinger v Kingsway Group Ltd (2009) 239 CLR 269

Commissioners for her Majesty's Revenue and Customs v The Investment Trust Companies (in liq) [2017] UKSC 29 (11 April 2017)

Crown Prosecution Services v Eastenders Group [2014] UKSC 26 (8 May 2014)

David Securities Pty Ltd v Commonwealth Bank of Australia (1992) 175 CLR 353

Deglman v Guaranty Trust of Canada [1954] SCR 725

Equuscorp Pty Ltd v Haxton (2012) 246 CLR 498

Farah Constructions Pty Ltd v Say-Dee Pty Ltd (2007) 230 CLR 89

Fibrosa Spolka Akcyjna v Fairbairn Lawson Combe Barbour Ltd [1943] AC 32

Ford v Perpetual Trustees Victoria Ltd (2009) 257 ALR 658

Garcia v National Australia Bank (1998) 194 CLR 395

Heperu Pty Ltd v Belle (2009) 258 ALR 727

Hill v Van Erp (1997) 188 CLR 159

Kakavas v Crown Melbourne Limited (2013) 250 CLR 392

Kiriri Cotton Co Ltd v Dewani [1960] AC 192

Lipkin Gorman v Karpnale Ltd [1991] 2 AC 548

Lowick Rose LLP (in liq) v Swynson [2017] UKSC 32 (11 April 2017)

Lumbers v W Cook Builders Pty Ltd (2008) 232 CLR 635

Martin v Pont [1993] 3 NZLR 25

Menelaou v Bank of Cyprus [2015] UKSC 66 (4 November 2015)

Moses v Macferlan (1760) 2 Burr 1005

National Mutual Life Association of Australasia Ltd v Walsh [1987] 8 NSWLR 585

Nepean District Tennis Association Inc v Penrith City Council (1988) 66 LGRA 440

Pavey & Matthews Pty Ltd v Paul (1987) 162 CLR 221

Pitt v Holt [2013] 2 AC 108

Public Trustee v Fraser [1987] 9 NSWLR 433

Roxborough v Rothmans of Pall Mall Australia Ltd (2001) 208 CLR 516

Royal Brunei Airlines Sdn Bhd v Tan [1995] 2 AC 378

Sempra Metals Ltd v Her Majesty's Commissioners for Inland Revenue [2007] UKHL 34 (18 July 2007)

Sinclair v Brougham [1914] AC 398

Slade's Case (1648) Style 138

Tame v New South Wales (2002) 211 CLR 317

Westdeutsche Landesbank Girozentrale v Islington LBC [1996] AC 669

Woolwich Equitable Building Society v Inland Revenue Commissioners [1993] AC 70

Legislation

Supreme Court Act 1970 (NSW)

Trade Practices Act 1974 (Cth)

Other

Edelman, James, 'Australian Challenges for the Law of Unjust Enrichment' (Speech delivered at the Summer School, University of Western Australia, 24 February 2012)

Virgo, Graham, 'Conscience or Unjust Enrichment?: The Emperor's Old Clothes: Australian Financial Services and Leasing Pty Ltd v Hills Industries Ltd' on *Opinions on High* (19 May 2014) <blogs.unimelb.edu.au/opinionsonhigh/2014/05/19/virgo-hills-industries/>

Unrepresentative swill? An unabashed defence of the Australian Senate

JULIAN MOSS

Abstract

This paper argues against the abolition of the Australian Senate. The Senate has endured decades of harsh criticism, exemplified by Paul Keating's label of 'unrepresentative swill'. Yet the institution continues to serve a number of useful purposes. Firstly, it enforces the strict horizontal separation of powers enshrined in the Constitution. Secondly, it serves as a state's house. That is, the Senate protects the rights of less populous states from the tyranny of the majority. In doing so, the chamber upholds Australia's federalist system. Finally, the Senate functions as a house of review to scrutinise hurried or reckless legislation in an auxiliary capacity. This paper concedes that the Senate in its current form is a flawed institution. As such, two packages of reforms are proposed. Firstly, senatorial terms may be reformed in order to reduce the incidence of rubber-stamping of legislation. Secondly, new states may be formed to reduce the population inequalities of Senate electorates while upholding the federalist system.

Introduction

The Senate is Australia's most disdained political institution. Paul Keating famously labelled the upper house as 'unrepresentative swill'.[1] Calls for reform are incessant.[2] In March 2016, Malcolm Turnbull succeeded in passing Senate electoral reforms designed to curb the electability of micro-party candidates.[3]

---

[1] Kieran Ricketts, 'The Collected Insults of Former PM Paul Keating', *ABC News* (online), 14 November 2014 <www.abc.net.au/news/2013-11-12/the-collected-insults-of-paul-keating/5071412>.

[2] See, eg, Terry Sweetman, 'We Need Senate Reform, Now More Than Ever', *The Courier Mail* (online), 15 November 2018 <myaccount.news.com.au/sites/couriermail/subscribe.html?sourceCode=CMWEB_WRE170_a_GGL&mode=premium&dest=https://www.couriermail.com.au/rendezview/we-need-senate-reform-now-more-than-ever/news-story/19320875da870f211cc47f2df4fbf21d&memtype=anonymous>.

[3] Eliza Borrello, 'Senate Reform: Electoral Laws Passed After Marathon Parliament Sitting', *ABC News* (online), 18 March 2016 <www.abc.net.au/news/2016-03-18/senate-electoral-reform-laws-passed/7258212>.

Recent research, however, predicts that minor parties will continue to hold a disproportionate influence over the Senate for the foreseeable future.[4] In retrospect, Turnbull's reforms appear largely ineffectual. The immense difficulty of reform begs the question: should the Senate simply be abolished? This paper argues that Australia's bicameral legislature should remain intact and that the Senate is a worthwhile institution, withstanding the shortcomings of the chamber. The first section of this paper shows that the Senate maintains a strict horizontal separation of powers. The second section asserts that the upper house protects the rights of states while upholding the federalist system. The third discusses the Senate's role as a house of review. Accordingly, the fourth section recommends several key reforms to restore integrity to the Senate in light of the shortcomings.

Horizontal separation of powers

A crucial argument for the retention of the Australian Senate is that a strict horizontal separation of powers should be preserved. Under the Australian convention of responsible government, the party with a majority in the House of Representatives forms the government.[5] The prime minister then advises the Governor-General to appoint ministers from those elected to parliament.[6] These ministers then implement laws, and by doing so effectively form the executive branch.[7] Thus, ministers simultaneously serve as members of the legislative and executive branches: they simply wear different hats. This contrasts with the political system of the United States, for example. The American president and his cabinet are elected separately from the members of congress.[8] Ergo, the distinction between the legislative and executive branches of government in the United States is clearer than that in Australia. Since the Commonwealth Constitution allots all states the same number of senators regardless of population,[9] the partisan makeup of the Senate

[4] Bill Browne, *'Get Used To It' Senate Projections, Autumn 2018* (The Australia Institute, 2018).

[5] Harry Evans (ed), *Odgers' Australian Senate Practice* (Australian Government Publishing Service, 7th ed, 1991) 11.

[6] Sir Percy Ernest Joske, *Australian Federal Government* (Butterworths, 3rd ed, 1976) 96.

[7] George Williams, Sean Brennan and Andrew Lynch, *Australian Constitutional Law and Theory: Commentary and Materials* (Federation Press, 6th ed, 2014) 352.

[8] Sir John A Cockburn, *Australian Federation* (Horace Marshall & Son, 1901) 45.

[9] *Australian Constitution* s 7.

can be and often is substantially different to that of the House of Representatives, and thus to the government of the day.[10]

Critically, the Senate demarcates the legislative and executive branches to prevent the rise of an autocratic regime. It does so by checking the power of the party in government.[11] Sir Percy Joske writes

> [t]he purpose of the upper house is to act as a curb, a brake or a stabilizer upon the lower house … which in absence of the upper chamber might wield despotic power.[12]

In addition, the Senate can hold the government to account by exercising its power to censure ministers.[13] This power is formidable, as it has led to ministerial resignations.[14] Consequently, the Senate's power to censure further segregates the legislative and the executive branches. This segregation maintains a strict horizontal separation of powers, and for this reason the Senate should not be abolished.

Although some have criticised the notion that the Constitution affirms a strict separation of powers, the Australian founding fathers expressly structured the Constitution to enunciate that affirmation. An example of this criticism may be found in *Dignan's Case*.[15] In that case, the Honourable Justice Evatt doubted the practicality of a strict separation of powers as it applies to delegated legislation. He stated:

> the full theory of 'Separation of Powers' cannot apply under our Constitution. Take the case of an enactment of the Commonwealth Parliament which gives to a subordinate authority other than the Executive, a power to make by-laws. To such an instance the theory of a hard and fast division and sub-division of powers between and among the three authorities of government cannot apply without absurd results.[16]

Denise Meyerson counters this argument by asserting that under the Constitution, the executive does not hold: 'the unfettered power to make law. One of the reasons why the framers of the Constitution vested

---

[10] See, eg, David Solomon, *Inside the Australian Parliament* (George Allen & Unwin, 1978) 183; Elaine Thompson, 'The Constitution and the Australian System of Limited Government, Responsible Government and Representative Democracy: Revisiting the Washminster Mutation' (2001) 24(3) *University of New South Wales Law Journal* 657, [8].

[11] Evans, above n 5, 13.

[12] Joske, above n 6, 73.

[13] Thompson, above n 10, [38].

[14] Thompson, above n 10, [39]–[40].

[15] *Victorian Stevedoring & General Contracting Co Pty Ltd v Dignan* ('*Dignan's Case*') (1931) 46 CLR 73.

[16] Ibid 119.

legislative, executive and judicial power in separate organs of government is that certain organs are more suitable for the performance of different governmental tasks.'[17] This is a crucial point.

While the drafters of the Constitution may not have foreseen legal problems arising from delegated legislation, they expressly formatted the Constitution to emphasise a strict horizontal separation of powers. The first three chapters of the Constitution are labelled 'The Parliament', 'The Executive Government', and 'The Judicature'.[18] Thus, from a constitutionalist perspective, this strict separation of powers should be preserved. The Senate occupies a unique position to examine delegated legislation 'not … enjoyed by the House of Representatives … because it is dominated by a disciplined majority supporting the government.'[19] Indeed, it has been primarily in the Senate that the executive government's use of delegated legislation has been effectively scrutinised.[20] This application of scrutiny further partitions the legislative and executive branches. Accordingly, the Senate should be retained in order to maintain the horizontal separation of powers.

A state's house

The Australian Senate should remain intact so that the rights of the states are protected and the federalist system is upheld. The Australian colonies would not have federated if it were not for the establishment of the Senate by the Constitution.[21] This is because the Senate advocates for the rights of individual states and the states as a whole under the Commonwealth.[22] Section 7 of the Constitution allots an equal number of senators to each state. Accordingly, the Senate protects the interests of the less populous states from being subjected to those of the populous states, whose representatives dominate the lower house.[23] In other words, the Senate acts as a bulwark against the tyranny of the majority. It is the only legislative body that gives a voice to the states, and therefore it is *the* keystone of Australian federalism. To abolish the Senate would be to topple the federalist system. In the words of Sir John Cockburn:

---

[17] Denise Meyerson, 'Rethinking the Constitutionality of Delegated Legislation' (2003) 11 *Australian Journal of Administrative Law* 45.

[18] *Australian Constitution* s 9.

[19] Evans, above n 5, 13.

[20] Evans, above n 5, 13.

[21] Joske, above n 6, 75.

[22] Joske, above n 6, 76.

[23] Evans, above n 5.

> [t]he whole principle of federation is to recognize the co-ordinate power of the population and of the States. There can be no federation if you give all the powers to the popular assembly.[24]

The argument of whether to abolish the Senate can be thought of as an argument over the advantages and disadvantages of replacing the federalist system with a more democratic unitary system. Critics of the Senate argue that since s 7 of the Constitution allots an equal number of senators to each state, electors from less populous states have a disproportionate influence over the upper house.[25] Elaine Thompson writes:

> because the Constitution gives equal representation to the States, and because the … States have different population[s], the Senate does not embody the notion of 'one vote, one value'. If that notion is seen as central to representative democracy, the Senate fails.[26]

The High Court in the *First Territory Senators Case*[27] demonstrated a preference for the ideological concept of representative democracy over the ideological concept of federalism.[28] However, is it true that the House of Representatives embodies the notion of 'one vote, one value'? In *McKinlay's Case*,[29] the High Court ruled that the Constitution does not mandate adherence to the notion of 'one vote, one value' when it comes to determining the number of electors in the electorates of the House of Representatives.[30] If these electorates are not required to be of equal population by the Constitution, it is difficult to argue that the House of Representatives embodies the notion. It is yet more difficult to argue that the Senate should be abolished for the reason that the chamber does not embody the notion of 'one vote, one value' and is therefore incompatible with the concept of representative democracy.

Even if one were to accept that the Senate is incompatible with representative democracy, there is a strong argument that the Constitution places more emphasis on federalism than representative democracy, despite the ruling in *First Territory Senators Case*. The High Court majority in *McGinty v Western Australia*[31] reaffirmed *McKinlay's Case*[32] and dismissed the argument that voter equality is implied in the

---

[24] *Australasian Federal Convention Debates*, Sydney, 16 March 1891, 383 (John A Cockburn).

[25] David Wood, 'The Senate, Federalism and Democracy' (1989) 17(2) *Melbourne University Law Review* 292, 294.

[26] Thompson, above n 10, [11].

[27] *Western Australia v Commonwealth ('First Territory Senators Case')* (1975) 134 CLR 201.

[28] Williams, Brennan and Lynch, above n 7, 694.

[29] *Attorney-General (Cth); Ex rel McKinlay v Commonwealth* (1975) 135 CLR 1.

[30] Williams, Brennan and Lynch, above n 7, 677.

[31] (1996) 186 CLR 140.

[32] *Attorney-General (Cth); Ex rel McKinlay v Commonwealth* (1975) 135 CLR 1.

Constitution.[33] More importantly, the majority held that such implications must be sourced from the text or structure of the Constitution and not from an amorphous concept such as representative democracy.[34] As per the Honourable Chief Justice Brennan: '[n]o implications can be drawn from the Constitution which is not based on the actual terms of the Constitution, or on its structure'.[35] Chapters V and VI are labelled 'The States' and 'New States' respectively, and s 7 provides for the establishment of the Senate as a states' house. Furthermore, the preamble to the Constitution states that the people of the colonies agreed to 'unite in one indissoluble Federal Commonwealth'. Indeed, the Constitution is replete with yet more references to federalism and federalist concepts.[36] Therefore, the concept of federalism is *clearly* endorsed by both the text and structure of the Constitution. The same cannot be said, however, for the concept of representative democracy.

Although after the turn of the century, the High Court has been willing to imply a qualified right to vote from the words 'directly chosen by the people' in ss 7 and 24 of the Constitution,[37] this implication of representative democracy is not as hardwired into the Constitution as federalism is. From a constitutional perspective, it is difficult to criticise the Senate for being insufficiently democratic. Accordingly, the Senate should not be abolished as a response to the criticism that the institution is incompatible with representative democracy.

A house of review

The Senate should be retained in order to serve as a house of review. A useful purpose of the Senate is the application of an additional layer of scrutiny over the legislative process.[38] This secondary opinion acts to prevent the assent of hurried or reckless legislation passed by the House of Representatives.[39] As discussed previously, the Senate occupies a unique position to monitor the executive government's use of delegated

[33] Williams, Brennan and Lynch, above n 7, 684.

[34] *McGinty v Western Australia* (1996) 186 CLR 140.

[35] Ibid [11].

[36] See, eg, *Australian Constitution* ss 73, 75, 92, 96, 100, and 101. See generally Bradley Selway, 'The Federation – What Makes It Work and What Should We Be Thinking About for the Future' (2001) 60(4) *Australian Journal of Public Administration* 116.

[37] *Roach v Electoral Commissioner* (2007) 233 CLR 162; *Rowe v Electoral Commissioner* (2010) 243 CLR 1.

[38] Evans, above n 5.

[39] Joske, above n 6, 31.

legislation.[40] This argument assumes a greater level of salience in the wake of the recent revival of federalism in the jurisprudence of the High Court in the *Pape*[41] and *Williams*[42] decisions. For example, in *Williams*,[43] the majority expressed concern that if an appropriation Act by itself would be sufficient to authorise expenditure by the executive, then senatorial scrutiny would be circumscribed.[44] This is because[45] appropriation Bills cannot be amended in the Senate.[46] In summary, the Constitution entrusts the Senate with the crucial role of acting as a legislative overseer. Accordingly, it should not be abolished.

While one may criticise the Senate's effectiveness as a house of review by asserting that it has become a 'party house', an obstructionist Senate may be thwarted by a double-dissolution election. Joske elucidates the 'party house' argument:

> when the government has a majority in the Senate … [it] functions as a house of review to little or no extent; and when opposition has the majority there, the Senate tends from a government point of view, to become obstructionist rather than to act as a body of review.[47]

The Constitution provides a mechanism for the Governor-General, on advice from the prime minister, to simultaneously dissolve the Senate and the House of Representatives in the event of a disagreement between the houses.[48] Therefore, voters can decide through election whether to return the government with a more favourable partisan composition of the Senate or conversely to vote the opposition into government. This clears parliamentary gridlock. Furthermore, it empowers the Australian people to decide whether the Senate has been unduly obstructing legislation or the government has hubristically passed pernicious legislation. Thus, the operation of s 57 solves the problem of a stonewalling, opposition-controlled Senate. However, it does not solve the problem of a government-controlled Senate rubber-stamping legislation. A solution to this problem, however, will be outlined in the following proposed reforms for the Australian Senate.

---

[40] Evans, above n 5, 12–13.

[41] *Pape v Commissioner of Taxation* (2009) 238 CLR 1.

[42] *Williams v Commonwealth of Australia* (2012) 248 CLR 156; *Williams v Commonwealth of Australia* [No 2] (2014) 252 CLR 416.

[43] *Williams v Commonwealth of Australia* (2012) 248 CLR 156.

[44] Shipra Chordia, Andrew Lynch and George Williams, 'Williams v the Commonwealth: Commonwealth Executive Power and Australian Federalism' (2013) 37(1) *Melbourne University Law Review* 189, 205.

[45] *Williams v Commonwealth of Australia* (2012) 248 CLR 156, 433 [60] (French CJ).

[46] *Australian Constitution* s 53.

[47] Joske, above n 6, 77.

[48] *Australian Constitution* s 57.

Proposals for reform

Reforms on terms

In regards to the problem of a government-controlled Senate rubber-stamping legislation, Senate reforms could be modelled on the proposed House of Lords Reform Bill 2011. This proposed bill provided, among other things, that elected members of the House of Lords would serve a single, non-renewable term of 15 years.[49] As these members would not be re-electable, threats issued by party leadership to reprimand them would be ineffectual.[50] Therefore, they would be politically unconstrained and at liberty to honestly scrutinise legislation.[51] If the Senate adopted these reforms, senators could refuse to rubber-stamp legislation without fear of repercussion from the government.

New states

Regarding the argument that the Senate gives disproportionate political power to the less populated states, new states could be formed as prescribed by s 124 of the Constitution. Populous states could be dissected, and less popular states could amalgamate by union. For example, Queensland could be split into two states: Northern Queensland and Southern Queensland.[52] The formation of new states could balance the population inequalities of the Senate electorates without toppling Australia's federalist system.

In addition, senators from smaller states would likely be more in touch with the specific needs of their constituents. Returning to the example of dividing Queensland into North and South, a hypothetical senator from Northern Queensland could focus on the issues affecting their largely rural constituents: namely balancing the competing interests of the cattle-grazing and coal-mining industries. Similarly, a hypothetical senator from Southern Queensland could focus on issues affecting their largely urbane constituents, such as improving transportation infrastructure between Brisbane and the Gold Coast.

Conclusion

The Australian Senate should not be abolished because it serves a number of useful purposes. It maintains a strict horizontal separation of powers. It is a State's house that upholds the federalist system. Furthermore,

---

[49] Joff Lelliot, 'What Australia Needs is a Genuine House of Review', *ABC News* (online), 3 January 2014 <www.abc.net.au/news/2014-01-03/lelliott-getting-rid-of-the-senate/5183256>.

[50] Lelliot, above n 49.

[51] Lelliot, above n 49.

[52] See, eg, Chris Berg, 'The democratic case for splitting Queensland in two', *ABC News* (online), 29 March 2016 <www.abc.net.au/news/2016-03-29/berg-the-democratic-case-for-splitting-queensland-in-two/7280330>.

it functions as a house of review. If reforms on term limits are implemented and new states are formed, problems plaguing the Senate can be remedied.

In the opinion of the author, the contemporary assault on the Senate is prompted by the difficulty that governments (both Labor and Liberal) face in advancing their legislative agendas. Additionally, the author speculates that intertwined veins of anti-Americanism and Anglophilia that flow through the Australian body politic promote hostility to federalism and affinity for a unitary system of government. The author humbly advises caution in tearing down constitutional checks and balances for the sake of political expediency and legislative convenience. History is replete with examples of this practice being accompanied with the rise of totalitarian regimes. Accordingly, we Australians should not be hasty in making fundamental changes to our constitutional structure. In the words of Friedrich Nietzsche: '[b]e careful, lest in casting out your demon you exorcise the best thing in you'.[53]

Acknowledgements

This article would have not come into existence were it not for Associate Professor Matthew Zagor, Professor and Head of School Anthony Connolly, and Associate Professor Ryan Goss. These three giants scooped me up and carried me on their shoulders. I profusely thank my father Robert Rex Moss for his emotional support and his intellectual tutelage. I thank my mother Annabel Julia Mead for challenging me to be the best version of myself. I thank my uncle Andrew Moss for putting a roof over my head when I was unable to secure accommodation in Canberra. I thank my dear friend Adrian Chak Hang Cheung for being there for me in my darkest hours.

I would also like to express deep gratitude to writer Robert Anson Heinlein, for inspiring my interest in political theory through his intrepid and thought-provoking works of fiction. I am deeply indebted to Dr Darryn Jensen, whose fervent passion for teaching and unwavering sense of duty exemplifies the Apollonian spirit. A special thanks goes out to the ANU College of Law support staff, in particular Ms Jayne Hardy, for tending to me when I experienced the misfortune of falling ill at the ANU Law Library. Finally, I am obligated to express gratitude to Mr Jay Wayne Jenkins, whose rousing and anthemic ballads bolstered my self-confidence when my spirits were low.

[53] Friedrich Nietzsche, *Thus Spoke Zarathustra: A Book for Everyone and No One* (RJ Hollingdale trans, Penguin Books, 1961) [trans of: *Also sprach Zarathustra: Ein Buch fur Alle und Keinan* (first published 1883–85)].

Bibliography

Articles/books/reports

Berg, Chris, 'The Democratic Case for Splitting Queensland in Two', *ABC News* (online), 29 March 2016 <www.abc.net.au/news/2016-03-29/berg-the-democratic-case-for-splitting-queensland-in-two/7280330>

Borrello, Eliza, 'Senate Reform: Electoral Laws Passed After Marathon Parliament Sitting', *ABC News* (online), 18 March 2016 <www.abc.net.au/news/2016-03-18/senate-electoral-reform-laws-passed/7258212>

Browne, Bill, *'Get Used To It' Senate Projections, Autumn 2018* (The Australia Institute, 2018)

Chordia, Shipra, Andrew Lynch and George Williams, 'Williams v the Commonwealth: Commonwealth Executive Power and Australian Federalism' (2013) 37(1) *Melbourne University Law Review* 189

Cockburn, Sir John A, *Australian Federation* (Horace Marshall & Son, 1901)

Evans, Harry (ed), *Odgers' Australian Senate Practice* (Australian Government Publishing Service, 7th ed, 1991)

Craven, Gregory (ed), *Australian Federation Towards the Second Century* (Melbourne University Pres, 1992)

Joske, Sir Percy Ernest, *Australian Federal Government* (Butterworths, 3rd ed, 1976)

Lelliot, Joff, 'What Australia Needs is a Genuine House of Review', *ABC News* (online), 3 January 2014 <www.abc.net.au/news/2014-01-03/lelliott-getting-rid-of-the-senate/5183256>

Meyerson, Denise, 'Rethinking the Constitutionality of Delegated Legislation' (2003) 11 *Australian Journal of Administrative Law* 45

Nietzsche, Friedrich, *Thus Spoke Zarathustra: A Book for Everyone and No One* (RJ Hollingdale trans, Penguin Books, 1961) [trans of: *Also sprach Zarathustra: Ein Buch fur Alle und Keinan* (first published 1883–85)]

Ricketts, Kieran, 'The Collected Insults of Former PM Paul Keating', *ABC News* (online), 14 November 2014 <www.abc.net.au/news/2013-11-12/the-collected-insults-of-paul-keating/5071412>

Sawyer, Geoffrey, *Australian Federalism in the Courts* (Melbourne University Press, 1967)

Selway, Bradley, 'The Federation – What Makes It Work and What Should We Be Thinking About for the Future' (2001) 60(4) *Australian Journal of Public Administration* 116

Solomon, David, *Inside the Australian Parliament* (George Allen & Unwin, 1978)

Sweetman, Terry, 'We Need Senate Reform, Now More Than Ever', *The Courier Mail* (Online), 15 November 2018 <myaccount.news.com.au/sites/couriermail/subscribe.html?sourceCode=CMWEB_WRE170_a_GGL&mode=premium&dest=https://www.couriermail.com.au/rendezview/we-need-senate-reform-now-more-than-ever/news-story/19320875da870f211cc47f2df4fbf21d&memtype=anonymous>

Thompson, Elaine, 'The Constitution and the Australian System of Limited Government, Responsible Government and Representative Democracy: Revisiting the Washminster Mutation' (2001) 24(3) *University of New South Wales Law Journal* 657

Vanstone, Amanda, 'Don't Bet Yet on a Double-Dissolution Election', *The Sydney Morning Herald* (online), 28 March 2016 <www.smh.com.au/comment/selfinterest-is-the-horse-running-hardest-20160326-gnrj2z.html>

Williams, George, Sean Brennan and Andrew Lynch, *Australian Constitutional Law and Theory: Commentary and Materials* (Federation Press, 6th ed, 2014)

Wood, David, 'The Senate, Federalism and Democracy' (1989) 17(2) *Melbourne University Law Review* 292

Cases

Attorney-General (Cth); Ex rel McKinlay v Commonwealth (1975) 135 CLR 1

McGinty v Western Australia (1996) 186 CLR 140

Pape v Commissioner of Taxation (2009) 238 CLR 1

Roach v Electoral Commissioner (2007) 233 CLR 162

Rowe v Electoral Commissioner (2010) 243 CLR 1.

Victorian Stevedoring & General Contracting Co Pty Ltd v Dignan ('Dignan's Case') (1931) 46 CLR 73.

Western Australia v Commonwealth ('First Territory Senators Case') (1975) 134 CLR 201

Williams v Commonwealth of Australia (2012) 248 CLR 156

Williams v Commonwealth of Australia [No 2] (2014) 252 CLR 416

Legislation

Australian Constitution

Other

Australasian Federal Convention Debates, Sydney, 16 March 1891, 383 (John A Cockburn)

Bird species richness and abundance: The effects of structural attributes, habitat complexity and tree diameter

YEE SENG TAY

Abstract

Structurally complex habitats influence species richness and abundance through the provision of resources and microhabitats. This study examines the effects of structural attributes, habitat complexity (expressed as habitat complexity score) and tree diameter (expressed as standard deviation of diameter at breast height), on bird species richness and abundance. Projected foliage cover and diameter at breast height of the five closest trees and bird richness and abundance were recorded at the 30 survey sites within the Kioloa Coastal Campus. Regression analyses were conducted between the structural attributes and bird richness and abundance. A statistically significant correlation (α=0.05) was found between habitat complexity score and bird richness and abundance. A similar relationship was observed with the standard deviation of diameter at breast height. Land managers can employ these findings to enhance habitat restoration and rehabilitation projects, and to monitor biodiversity outcomes.

Introduction

Over the last two hundred years, 29 native Australian bird species have gone extinct (Szabo, Khwaja, Garnett, & Butchart, 2012) and this figure is forecasted to reach 39 by 2038 unless biodiversity and land management practices improve drastically (Geyle et al., 2018). The concept of habitat complexity, or the structural attributes of an environment (Smith, Johnston, & Clark, 2014), was first introduced in the classical work 'On Bird Species Diversity' by MacArthur and MacArthur (1961), who discovered a positive relationship between foliage height diversity and bird species diversity. Since then, the study of habitat complexity and its effects has proliferated the field of ecological science, with similar correlations being observed in the terrestrial (Gardner, Cabido, Valladares, & Diaz, 1995; Lawton, 1983) and marine domains (Kelaher & Castilla, 2005; Luckhurst & Luckhurst, 1978). Researchers generally seek to identify structural attributes that affect species richness and abundance. They hypothesise that complex habitats offer more potential niches and reduce the possibility of niche overlap as compared to structurally simpler habitats (Klopfer & MacArthur, 1960, 1961).

Integrating the ecological niche theory popularised by Hutchinson (1957) with birds, one can deduce that in structurally complex habitats, birds will have access to a variety of foraging sites from which specialisation can occur, enabling more species to be added into the existing assemblage (Karr, 1990). For example, the presence of ground litter and coarse woody debris in complex landscapes creates the microhabitats required to support the invertebrate community, an important source of food for many bird species (Nilsson, 1979). Apart from food, complex habitats increase the availability of nesting, perching and roosting sites, which are fundamental to birds, especially tree-hollow nesters. For example, Gibbons, Lindenmayer, Barry, and Tanton (2000) established that structural attributes such as tree diameter are associated with the occurrence of tree hollows and this may implicate hollow nesters such as the threatened superb parrot (*Polytelis swainsonii*) (Manning, Gibbons, Fischer, Oliver, & Lindenmayer, 2013).

Tree diameter is also an important stand structural attribute that communicates the health of habitats, ecological functions and the process of ecological succession (Spies & Franklin, 1991), which may affect bird richness and abundance. Thus, complex habitats can increase the availability of resources, decreasing competition and resulting in increased bird richness and abundance (Seddon, Briggs, & Doyle, 2001, 2003).

This paper aims to contribute to existing land management practices and bird conservation efforts by identifying key relationships between structural attributes and bird richness and abundance. Stakeholders can employ these findings to facilitate land restoration projects, habitat rehabilitation and monitoring. The study methods and findings discussed below explore the effects of structural attributes on bird species richness and abundance. In line with the previous observations, it is hypothesised that bird species richness and abundance increase as habitats become more complex.

Methods

Study area

The study was carried out at the Kioloa Coastal Campus (KCC) of ANU. KCC is located in the south-eastern coastal district of Shoalhaven, New South Wales (35°32'S; 150°22'E). It encompasses an area of 348 hectares, which borders the foothills of the Murramarang Ranges to the west, the mean high-water mark of the beaches to the east, private lands to the north and Crown land such as the Murramarang National Park to the south (Figure 1).

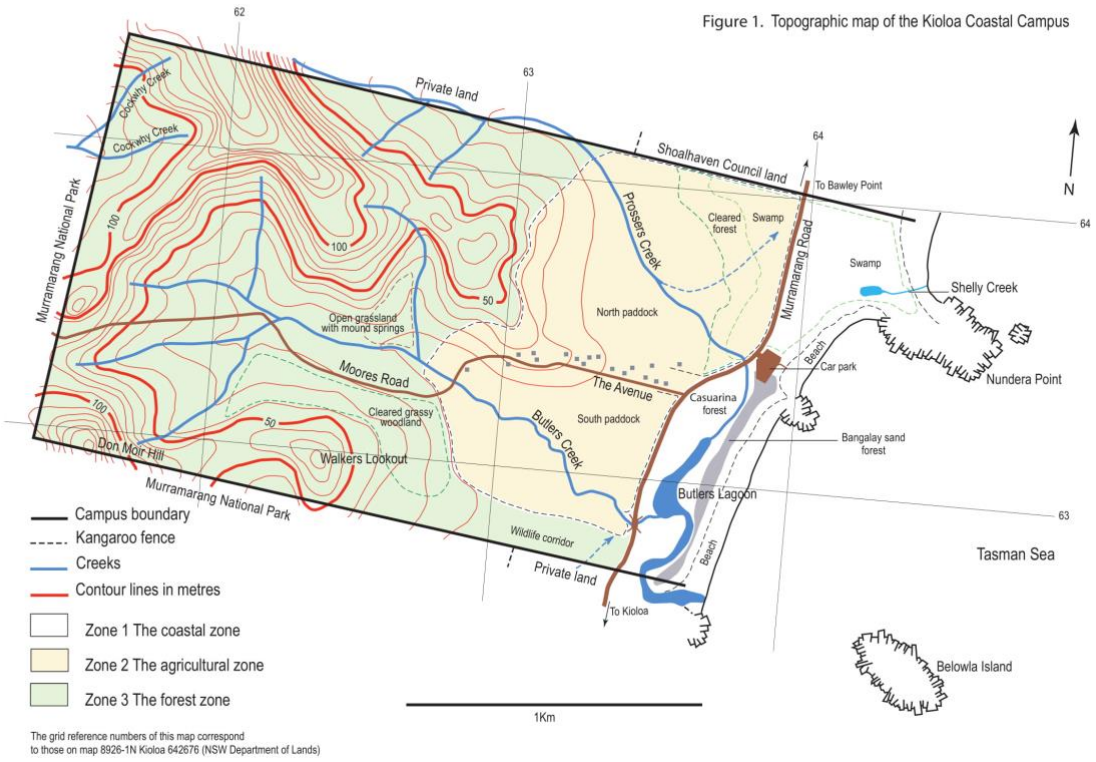

Figure 1: Annotated Map of KCC.

Source: ANU Facilities & Services Division, 2017.

KCC receives a mean annual rainfall of 1,110 mm and experiences mean annual temperature of 13.2°C to 20.6°C (BOM, 2017). Elevation within KCC ranges from 0 to 100 metres above sea level. KCC can be categorised into three major vegetation zones, coastal, agricultural and forest (Figure 1), which reflects on the land use history of the property (Caton, 2007). This study focuses on agricultural and forest zones. The former comprises of cleared pastures and remnant vegetation along creeks and is occupied by tussock grass (*Poa sp.*) and grasses such as ryegrass (*Lolium multiflorum*) and kikuyu (*Pennisetum clandestinum*). The latter can be further classified into regrowth and old growth areas.

The regrowth area comprises of cleared grassy woodlands, with spotted gums (*Corymbia maculata*), prickly beard-heath (*Leucopogon juniperinus*) and orange thorn (*Citriobatus pauciflorus*). The old growth area comprises of southern lowland wet sclerophyll forest and southern warm temperate rainforest. Spotted gums (*C. maculata*), blackbutt (*Eucalyptus pilularis*), bangalays (*E. botryoides*) and stringybarks (*E. spp.*) occupy the sclerophyll forest and bolwarras (*Eupomatia laurina*), lilly pillies (*Syzygium smithii*) and tree ferns (*Dicksonia antarctica* and *Cyathea australis*) occupy the rainforest.

Bird survey

This study employed a total of 30 survey sites within the KCC and the sites were allocated equally among the agricultural (pasture) and forest (regrowth and old growth) zones (Figure 2).

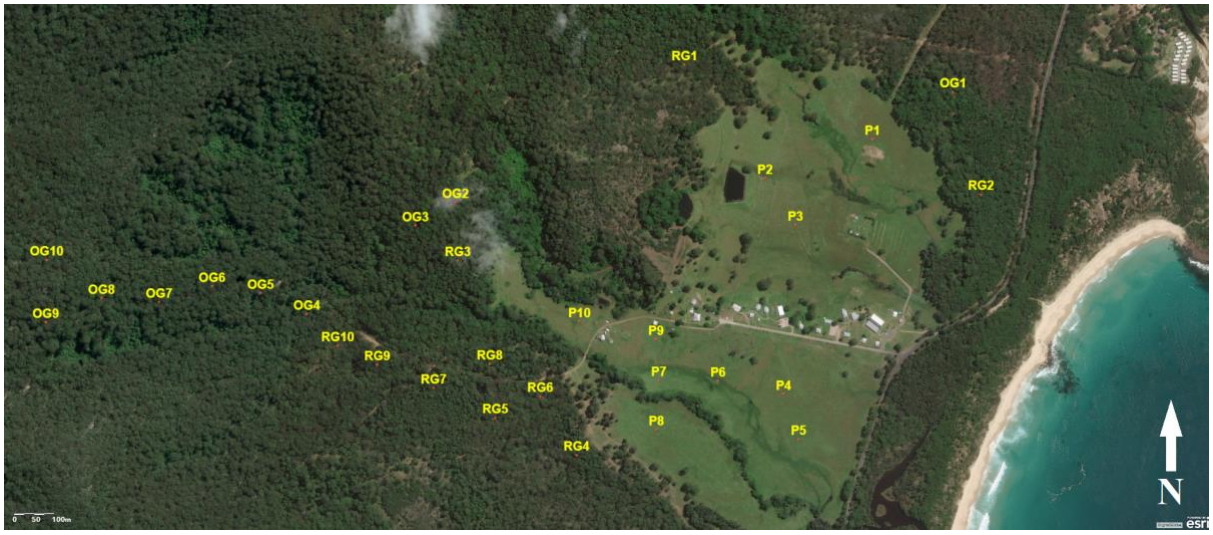

Figure 2: Satellite image of KCC and the location of the 30 survey sites.

Source: Esri, 2017.

Bird surveys were conducted in the early hours of the day following sunrise, between 0700H to 1000H, as bird activity peaks during this period and gradually declines as the morning progresses, as suggested by Robbins (1981). The 30 sites were equally divided among six groups and a trained expert/ornithologist was assigned to each group. The bird surveys were conducted using the single visit point count method as outlined by Laiolo (2002), where all birds seen or heard within a 50-metre radius of the site were recorded. Overflying birds that were passing by were not considered to be using the site and they were not recorded in our study, as suggested by Laiolo (2002). Each site was sampled for five minutes in order to achieve maximum efficiency when recording the number of species per unit effort, as suggested by WP Smith et al. (1993). The total number of birds seen or heard was recorded as bird abundance and the total number of different bird species observed was recorded as species richness.

Vegetation structure survey

At each site, the dominant species at each level of the stratum (upper, middle and ground) were recorded. The diameter at breast height (DBH) of the five nearest trees was measured using a diameter tape. Breast height is defined as 1.30 metres above the ground (Brack & Wood, 1998). Subsequently, a 100-metre transect was established across the site using the measuring tape. The bearing of the transect was chosen in a direction that was representative of the site's vegetation structure. The point intercept method, which involves the densitometer, was used to determine the presence or absence of foliage cover in the upper and mid stratum, as outlined by Hnatiuk, Thackway, and Walker (2010).

The presence or absence of foliage cover (ground) was conducted in a rudimentary manner but it adhered closely to the point intercept method, where the surveyor's toe on the right boot was used as the intercept point. Foliage cover (ground) consisted of grass tussocks (>10 cm in height), fine herbaceous (<10 cm in height), litter (detached plant material), rock, bare ground and coarse woody debris (>5 cm in diameter). A total of 50 observations were made along the transect, at intervals of 2.00 metres and the projected foliage cover was obtained for each component by dividing the sum of present counts by 50 (total number of observations) and multiplying it by 100 (conversion to percentages).

Data analysis

With respect to the research question and hypothesis, the dependent variables in this study are defined as the total number of bird species (richness) and the total number of birds (abundance). The independent variables are defined as the standard deviation of DBH of the five closest trees (SD.DBH) and habitat complexity, in terms of the habitat complexity score (HCS). Microsoft Excel was used to perform the bivariate regression analysis between the dependent and independent variables, to determine if a statistically significant relationship exists between the structural attributes and bird species richness and abundance.

Tree DBH data collected at each site was processed through Microsoft Excel using the standard deviation function to obtain SD.DBH. The data was quantified and analysed in terms of SD.DBH as it reflected the logging and disturbance history of the site (McElhinny, Gibbons, & Brack, 2006) and serves as an indicative measure of biological diversity at the site (Buongiorno, Dahir, Lu, & Lin, 1994). The HCS for each site was obtained by inputting the projecting foliage cover for each component into the scoring matrix below (Table 1).

Table 1: The components of the habitat complexity score and method of scoring

	Score 0	**Score 1**	**Score 2**	**Score 3**	**Comments**
Component	0–10%	10–20%	20–50%	>50%	
Upper stratum (% cover)					
Mid stratum (% cover)					
Ground cover (% cover)*					
Litter cover (% cover)					
Component	0–5%	5–10%	10–15%	>15%	
Rock (% cover)					
Coarse wood debris (% cover)					

Bare ground (% cover)					
Litter cover (% cover)					
Total					Sum** =

*Comprises of projected foliage cover data of grass tussocks and fine herbaceous.

**Refers to the sum of the total score for each scoring column (0–3).

Source: Table created by the author, based on Catling and Burt (1995) and Watson, Freudenberger, and Paull (2001).

The scoring matrix is based on Catling and Burt (1995) and Watson, Freudenberger, and Paull (2001) and has been adapted for this study. The modifications consisted of assigning new per cent cover values for each score under rocks, coarse woody debris and bare ground, to capture a greater spread of the data collected. Additionally, the ground cover component in the matrix includes both grass tussock and fine herbaceous projected foliage cover data.

Results

Bird species richness and abundance

A total of 645 individual birds belonging to 44 species were recorded during the bird survey. Bird species richness ranged 0–9 species at the pasture sites (2.90 ± 1.02; mean ± standard error), 7–18 at the regrowth sites (12.50 ± 1.01) and 10–16 at the old growth sites (12.70 ± 0.68). Bird species abundance ranged 0–22 at the pasture sites (7.40 ± 2.65), 15–55 at regrowth sites (30.60 ± 3.39) and 22–38 at the old growth sites (26.5 ± 1.74).

Vegetation structure

Tree DBH ranged from 5.00 cm to 72.00 cm at the regrowth sites (28.00 cm ± 2.57) and from 3.50 cm to 350.00 cm at old growth sites (63.34 cm ± 7.50). No trees were detected at the pasture sites. The range and mean projected foliage cover for each habitat type is presented in Table 2.

Table 2: Range and mean of projected foliage cover for each component for each habitat type

Habitat type	**Upper stratum, %**	**Mid stratum, %**	**Ground cover, %**	**Litter cover, %**	**Rock, %**	**Course woody debris, %**	**Bare ground, %**
Pasture	0	0	84%–100% 87% ± 8.76	8%–16% 4% ± 1.91	0%–2% 0% ±	0%–2% 0% ±	0%–10% 1% ± 0.99
Regrowth	54%–96% 74% ± 3.79	6%–48% 25% ± 3.70	4%–48% 31% ± 4.84	30%–90% 65% ± 5.82	0%–26% 5% ± 2.59	0%–6% 3% ± 0.82	0%–2% 1% ± 0.33

Old growth	22%–92% 67% ± 6.08	14%–90% 42% ± 7.31	0%–64% 17% ± 6.72	2%–94% 67% ± 9.13	0%–14% 2% ± 1.51	0%–18% 8% ± 1.81	0%–14% 3% ± 1.34

Source: Author's summary of fieldwork data.

The dominant species at the ground stratum was identified as kikuyu (*P. clandestinum*) for pasture sites, cogon grass (*Imperata cylindrica*) for regrowth sites and austral bracken (*Pteridium esculentum*) for old growth sites. The dominant species at the middle stratum was identified as native blackthorn (*Bursaria spinosa*) at the regrowth sites and wild yellow jasmine (*Pittosporum revolutum*) for old growth sites. The dominant species at the upper stratum was identified as spotted gums (*C. maculata*) for both regrowth and old growth sites.

Bivariate regression analysis, richness

A significant positive correlation was observed ($R^2 = 0.62$, $p = <0.001$) between HCS and the total number of bird species (richness) (Table 3). A similar relationship was observed ($R^2 = 0.39$, $p = <0.001$) between SD.DBH and the total number of bird species (richness) (Table 3).

Table 3: Results from linear regression analysis for predicting bird species richness

Independent variable	Coefficient	Standard error	Confidence interval (a=0.05)	R^2	F	t-value	p-value
HCS	1.21	0.18	0.84–1.57	0.62	46.08	6.79	<0.001
SD.DBH	0.13	0.03	0.07–0.20	0.39	17.81	4.22	<0.001

Source: Author's summary of fieldwork data.

Both relationships are illustrated by the graphs in Figures 3 and 4, respectively.

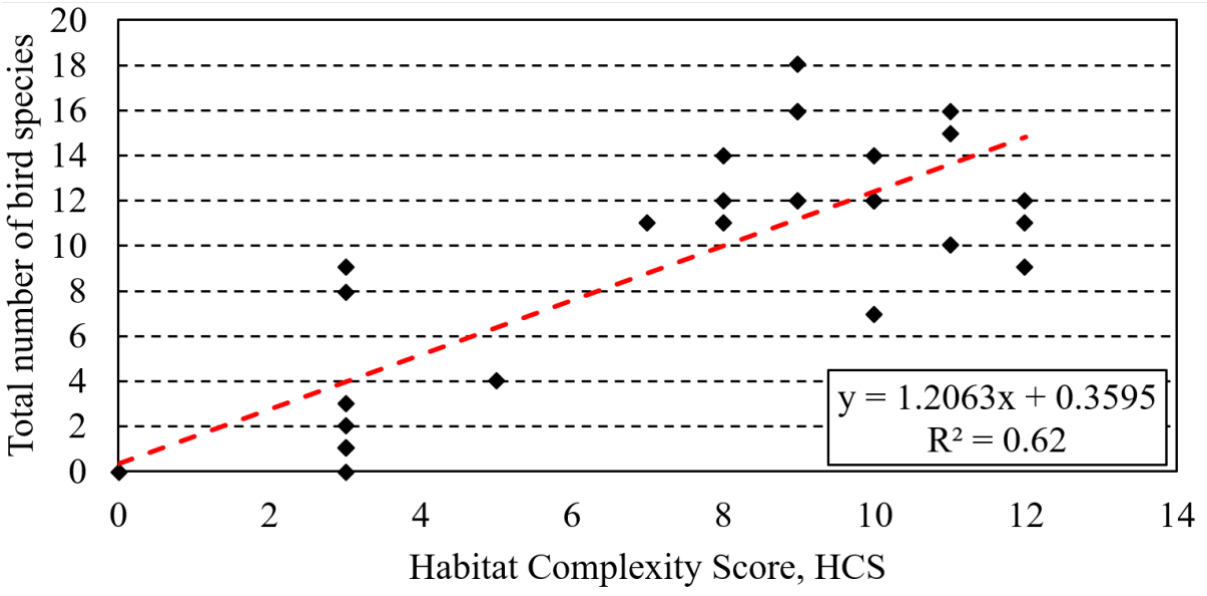

Figure 3: Bivariate linear regression analysis of total number of bird species (richness) with HCS.

Source: Author's summary of fieldwork data.

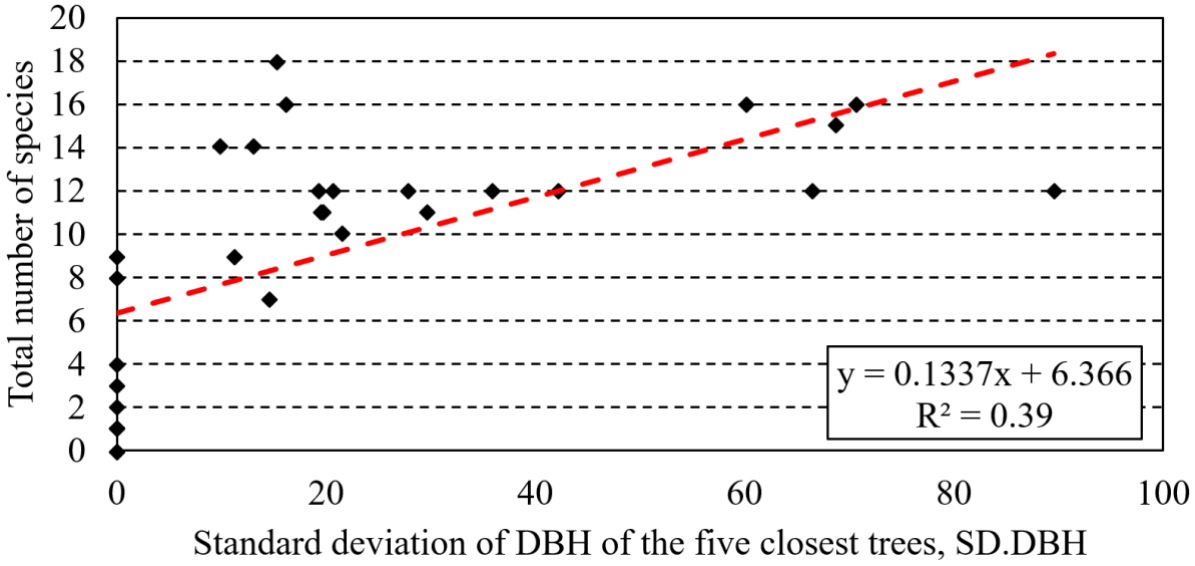

Figure 4: Bivariate linear regression analysis of the total number of bird species (richness) with SD.DBH.

Source: Author's summary of fieldwork data.

Bivariate regression analysis, abundance

A significant positive correlation was observed ($R^2 = 0.48$, $p = <0.001$) between HCS and the total number of birds (abundance) (Table 4). A similar relationship was observed ($R^2 = 0.14$, $p = 0.043$) between SD.DBH and the total number of birds (abundance) (Table 4).

Table 4: Results from linear regression analysis for predicting bird species abundance

Independent variable	Coefficient	Standard error	Confidence interval (a=0.05)	R^2	F	t-value	p-value
HCS	2.57	0.50	1.54–3.60	0.48	26.16	5.11	<0.001
SD.DBH	0.72	0.34	0.002–1.41	0.14	4.50	2.12	0.0043

Source: Author's summary of fieldwork data.

Both relationships are illustrated by the graphs in Figures 5 and 6, respectively.

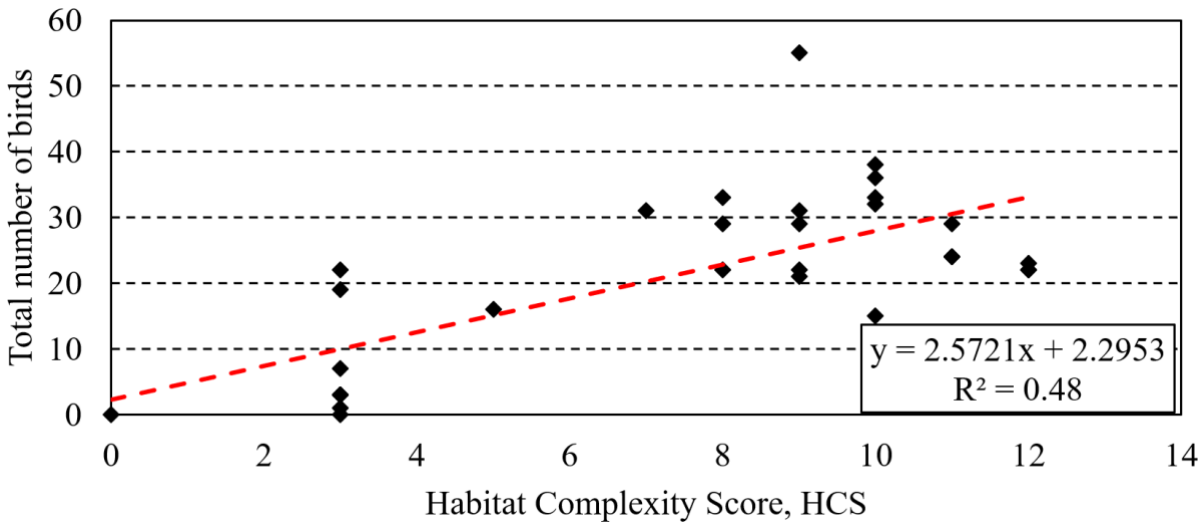

Figure 5: Bivariate linear regression analysis of the total number of birds (abundance) with the HCS.

Source: Author's summary of fieldwork data.

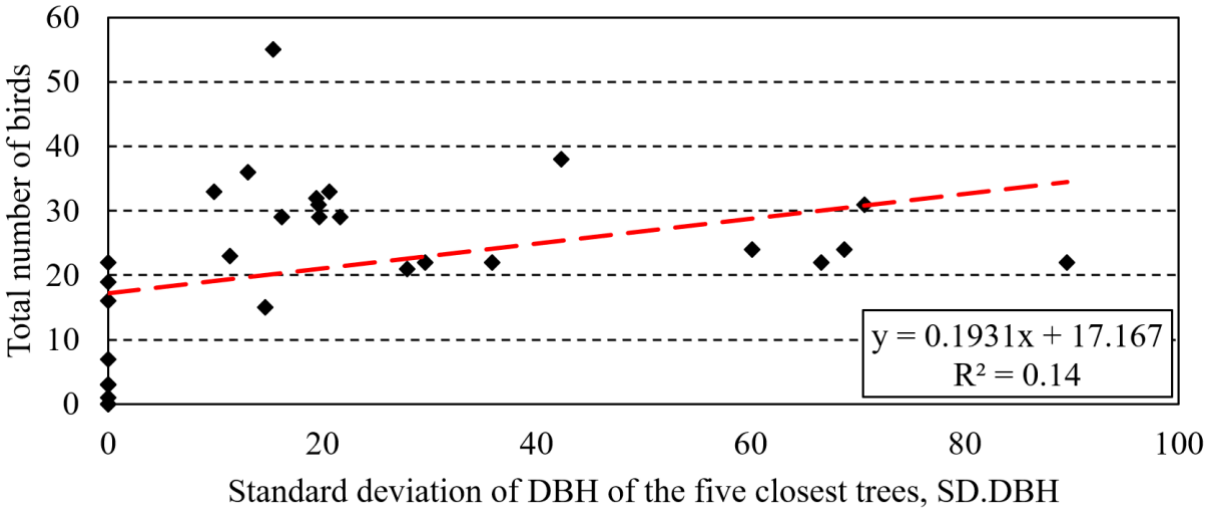

Figure 6: Bivariate linear regression analysis of the total number of birds (abundance) with the variation in DBH of the five closest trees.

Source: Author's summary of fieldwork data.

Discussion

This study investigated the relationship between structural attributes and bird species richness and abundance. Results supported the hypothesis that bird richness and abundance increase as habitats become more complex. That is, a statistically significant correlation was found between HCS and bird richness and abundance, and a similar relationship was observed with SD.DBH.

Bird richness and abundance with HCS

The regression analysis yielded a significant correlation between habitat complexity, expressed in terms of HCS, and bird species richness and abundance. Based on the results, 62 per cent and 48 per cent of the variance in richness and abundance respectively can be accounted for by habitat complexity. This implies that bird species richness and abundance increase in proportion to habitat complexity. This relationship can be explained by structural attributes found within each survey site. The absence of trees within pasture sites effectively limits the richness and abundance of birds as there are no incentives for hollow-nesting birds to be present, since the habitat offers no nesting sites (Newton, 1994).

The absence of trees also meant that the in-situ production of coarse woody debris would be non-existent and litter would be limited to wind transportation and ground production, preventing the formation of suitable microhabitats that would allow invertebrates to thrive (Berg et al., 1994), disincentivising insectivorous birds. The absence of upper and mid-canopy cover in the pasture sites may generate unsuitable temperature conditions, negatively affecting bird species richness (Pearman, 2002). Importantly, the low structural complexity of the pasture sites, relative to regrowth and old growth, limits the availability of ecological niches for birds to specialise in, increasing competition and restricting the realised niche to the most adept (Keller & Lloyd, 1994).

Conversely, in the regrowth and old growth sites, the increased structural complexity facilitates the division of habitat resources among different bird species (Hurlbert, 2004). This reduces competition and allows for niche specialisation, enabling greater richness (Van Den Meersschaut & Vandekerkhove, 2000).

Bird richness and abundance with SD.DBH

Based on the regression analysis, 39 per cent and 14 per cent of the variance in richness and abundance respectively can be accounted for by SD.DBH. This implies that bird species richness and abundance increase in habitats where both old (larger DBH value) and young (smaller DBH value) trees coexist. Older trees are more likely to have experienced stochastic heat stress events such as bushfires, which promotes the process of hollow formation in addition to natural formation with age (Adkins, 2006). Younger trees are also not exempted from such events and can develop hollows as a result of fire (Gibbons et al., 2000). Thus, increased hollow abundance can be expected in trees of larger diameter (Gibbons et al., 2000; Lindenmayer, Cunningham, Donnelly, Tanton, & Nix, 1993).

The presence of hollows does not equate to greater bird species richness or abundance as birds are not the only fauna that use hollows and birds prefer hollows with entrances similar to the size of their body width (Goldingay, 2009). The resulting abundance in tree hollows of varying sizes due to the presence of older and younger trees increases the likelihood of hollow suitability for birds. This reduces

competition among birds for nesting sites, increasing breeding density and promoting greater bird richness and abundance (Newton, 1994).

Most importantly, SD.DBH reflects on the assortment of microhabitats present at the stand (Acker, Sabin, Ganio, & McKee, 1998; Van Den Meersschaut & Vandekerkhove, 2000). The variation observed in the old growth forests, where both young and old trees were observed to coexist, also suggests a rich habitat with a variety of niches and resources (Temple, Mossman, & Ambuel, 1979). This allows for specialisation and reduces niche overlap, decreasing competition and relaxing the realised niche.

Limitation of study

Regarding the experimental design of this study, the location of some survey sites (Figure 2) appears to have been placed at regular intervals along Moores Road for the purpose of accessibility. This may introduce additional variables into our data, as the road is a form of human interference and has changed the landscape around it. The survey sites can be improved through randomised allocation within the forest zone and a buffer can be established to prevent a site from being allocated within a certain distance from the road.

Additionally, the collection of the projected foliage cover data was conducted in a subjective manner, where some groups established transect in a bearing that prioritised accessibility over the accurate representation of the site. This issue can be mitigated by ensuring that surveyors are dressed and geared properly for the job, to reduce and remove any hesitation regarding fieldwork. Geographic Informational Systems (GIS) can also be used to remotely sense certain components of the projected foliage cover, to improve the replicability and the precision of data collected.

Bird species richness and abundance may not have been captured accurately at some of the survey sites due to the way surveys were conducted. Bird activity declines as the morning progresses, as suggested by Robbins (1981) and this would mean that the bird surveys conducted by groups for the last few sites would reflect a lower bird species richness and abundance. This issue can be resolved by ensuring that the survey team is prepared before sunrise so that survey can be conducted within the golden hour of after sunrise. The number of bird surveys conducted can also be increased to reduce the effect of natural variation, increasing precision, and the flow of the sites to survey can be randomised to reduce the effect imposed by the time of day on the bird data.

Implications

This study has established the effects of structural attributes, habitat complexity and tree diameter on bird species richness and abundance. Land managers and conservationists can employ the findings of this study to better manage their landscape, such as retaining mature hollow-bearing trees during timber extraction or prescribing regular fires to promote hollow formation. Bird habitats can also be

deliberately restored or enhanced by tweaking the settings of structural attribute components, such as coarse woody debris and litter cover, which support the invertebrates that most birds feed on.

Alternatively, trees with differing growth rates can be planted to create variability in tree diameter, which enhances microhabitat diversity and resource availability. Importantly, this study highlights the importance of an evidence-based management approach in promoting positive biodiversity outcomes, and the value of future research in stemming the mass decline of Australia's biodiversity.

Acknowledgements

I would like to express my deepest gratitude to Associate Professor Philip Gibbons of the Fenner School of Environment and Society for his patient guidance, enthusiastic encouragement, and constructive criticisms during the development of this research work.

References

Acker, S, Sabin, T, Ganio, L, & McKee, W. (1998). Development of old-growth structure and timber volume growth trends in maturing Douglas-fir stands. *Forest Ecology and Management, 104*(1), 265–80. doi.org/10.1016/S0378-1127(97)00249-1.

Adkins, MF. (2006). A burning issue: Using fire to accelerate tree hollow formation in Eucalyptus species. *Australian Forestry, 69*(2), 107–13. doi.org/10.1080/00049158.2006.10676236.

ANU Facilities & Services Division. (2017). Topographic map of the Kioloa Coastal Campus. *Kioloa Coastal Campus*. Retrieved from services.anu.edu.au/files/business-unit/topographic-map.pdf.

Berg, Å, Ehnström, B, Gustafsson, L, Hallingbäck, T, Jonsell, M, & Weslien, J. (1994). Threatened plant, animal, and fungus species in Swedish forests: Distribution and habitat associations. *Conservation biology, 8*(3), 718–31. doi.org/10.1046/j.1523-1739.1994.08030718.x.

BOM (Bureau of Meteorology). (2017). Summary statistics Batemans Bay (Catalina Country Club). Retrieved from bom.gov.au/climate/averages/tables/cw_069134.shtml.

Brack, C, & Wood, G. (1998). Forest mensuration–Measuring trees, stands and forests for effective forest management [Online tool and guidebook]. ISBN 0 7315 2413 6. Retrieved from fennerschool-associated.anu.edu.au/mensuration/BrackandWood1998/MENSHOME.HTM.

Buongiorno, J, Dahir, S, Lu, H-C, & Lin, C-R. (1994). Tree size diversity and economic returns in uneven-aged forest stands. *Forest Science, 40*(1), 83–103.

Catling, P, & Burt, R. (1995). Studies of the ground-dwelling mammals of eucalypt forests in south-eastern New South Wales: The effect of habitat variables on distribution and abundance. *Wildlife Research, 22*(3), 271–88. doi.org/10.1071/WR9950271.

Caton, JM. (2007). *The biodiversity of the Kioloa coastal campus of the Australian National University* (unpublished report). Retrieved from

www.researchgate.net/profile/Klaus_Henle/publication/275123874_The_herpetofauna_of_Kioloa_New_South_Wales_Baseline_observational_data_collected_30_years_ago_and_inspired_by_RE_Barwick/links/5835815308ae138f1c113b64/The-herpetofauna-of-Kioloa-New-South-Wales-Baseline-observational-data-collected-30-years-ago-and-inspired-by-RE-Barwick.pdf.

Esri. (2017). Topographic basemap of Kioloa Coastal Campus, scale not given. *World Topographic Map.* Retrieved from www.arcgis.com/home/webmap/viewer.html?webmap=a20f13a1890a4e72ab1a9accf44f38e8.

Gardner, SM, Cabido, MR, Valladares, GR, & Diaz, S. (1995). The influence of habitat structure on arthropod diversity in Argentine semi-arid Chaco forest. *Journal of Vegetation Science, 6*(3), 349–56. doi.org/10.2307/3236234.

Geyle, HM, Woinarski, JC, Baker, GB, Dickman, CR, Dutson, G, Fisher, DO, … Kutt, A. (2018). Quantifying extinction risk and forecasting the number of impending Australian bird and mammal extinctions. *Pacific Conservation Biology, 24*(2), 157–67. doi.org/10.1071/PC18006.

Gibbons, P, Lindenmayer, D, Barry, SC, & Tanton, M. (2000). Hollow formation in eucalypts from temperate forests in southeastern Australia. *Pacific Conservation Biology, 6*(3), 218. doi.org/10.1071/PC000217.

Goldingay, RL. (2009). Characteristics of tree hollows used by Australian birds and bats. *Wildlife Research, 36*(5), 394–09. doi.org/10.1071/WR08172.

Hnatiuk, RJ, Thackway, R, & Walker, J. (2010). Vegetation. In JG Speight & RF Isbell (eds), *Australian Soil and Land Survey Field Handbook* (3rd ed., pp. 239). Collingwood: CSIRO Publishing.

Hurlbert, AH. (2004). Species–energy relationships and habitat complexity in bird communities. *Ecology Letters, 7*(8), 714–20. doi.org/10.1111/j.1461-0248.2004.00630.x.

Hutchinson, GE. (1957). Concluding remarks. *Cold Spring Harbor symposium on quantitative biology, 22*, 415–27. doi.org/10.1101/SQB.1957.022.01.039.

Karr, JR. (1990). Interactions between forest birds and their habitats: a comparative synthesis. *Biogeography and ecology of forest bird communities*, 379–86.

Kelaher, BP, & Castilla, JC. (2005). Habitat characteristics influence macrofaunal communities in coralline turf more than mesoscale coastal upwelling on the coast of Northern Chile. *Estuarine, Coastal and Shelf Science, 63*(1-2), 155–65. doi.org/10.1016/j.ecss.2004.10.017.

Keller, EF, & Lloyd, EA. (1994). *Keywords in evolutionary biology*: Harvard University Press.

Klopfer, PH, & MacArthur, RH. (1960). Niche size and faunal diversity. *The American Naturalist, 94*(877), 293–300. doi.org/10.1086/282130.

Klopfer, PH, & MacArthur, RH. (1961). On the causes of tropical species diversity: Niche overlap. *The American Naturalist, 95*(883), 223–26. doi.org/10.1086/282179.

Laiolo, P. (2002). Effects of habitat structure, floral composition and diversity on a forest bird community in north-western Italy. *Folia Zoologica, 51*(2), 121–28.

Lawton, J. (1983). Plant architecture and the diversity of phytophagous insects. *Annual Review of Entomology, 28*(1), 23–39. doi.org/10.1146/annurev.en.28.010183.000323.

Lindenmayer, D, Cunningham, R, Donnelly, C, Tanton, M, & Nix, H. (1993). The abundance and development of cavities in Eucalyptus trees: A case study in the montane forests of Victoria, southeastern Australia. *Forest Ecology and Management, 60*(1-2), 77–104. doi.org/10.1016/0378-1127(93)90024-H.

Luckhurst, B, & Luckhurst, K. (1978). Analysis of the influence of substrate variables on coral reef fish communities. *Marine Biology, 49*(4), 317–23. doi.org/10.1007/BF00455026.

MacArthur, RH, & MacArthur, JW. (1961). On bird species diversity. *Ecology, 42*(3), 594–98. doi.org/10.2307/1932254.

Manning, A, Gibbons, P, Fischer, J, Oliver, D, & Lindenmayer, D. (2013). Hollow futures? Tree decline, lag effects and hollow-dependent species. *Animal Conservation, 16*(4), 395–403. doi.org/10.1111/acv.12006.

McElhinny, C, Gibbons, P, & Brack, C. (2006). An objective and quantitative methodology for constructing an index of stand structural complexity. *Forest Ecology and Management, 235*(1), 54–71. doi.org/10.1016/j.foreco.2006.07.024.

Newton, I. (1994). The role of nest sites in limiting the numbers of hole-nesting birds: a review. *Biological Conservation, 70*(3), 265–76. doi.org/10.1016/0006-3207(94)90172-4.

Nilsson, SG. (1979). Density and species richness of some forest bird communities in South Sweden. *Oikos*, 392–401. doi.org/10.2307/3544327.

Pearman, PB. (2002). The scale of community structure: Habitat variation and avian guilds in tropical forest understory. *Ecological Monographs, 72*(1), 19–39. doi.org/10.1890/0012-9615(2002)072[0019:TSOCSH]2.0.CO;2.

Robbins, CS. (1981). Effect of time of day on bird activity. *Studies in Avian Biology, 6*, 275–86.

Seddon, J, Briggs, S, & Doyle, S. (2001). *Birds in woodland remnants of the central wheat/sheep belt of New South Wales: Report on NHT Project, predicting biodiversity of woodland remnants for on-ground conservation* (Report). NSW National Parks and Wildlife Service.

Seddon, J, Briggs, S, & Doyle, S. (2003). Relationships between bird species and characteristics of woodland remnants in central New South Wales. *Pacific Conservation Biology, 9*(2), 95–119. doi.org/10.1071/PC030095

Smith, RS, Johnston, EL, & Clark, GF. (2014). The role of habitat complexity in community development is mediated by resource availability. *PloS one, 9*(7), e102920.

Smith, WP, Twedt, DJ, Wiedenfeld, DA, Hamel, PB, Ford, RP, & Cooper, RJ. (1993). *Point counts of birds in bottomland hardwood forests of the Mississippi Alluvial Valley: Duration, minimum sample size, and points versus visits* (Research Paper SO-274). New Orleans, Louisiana: US Department of Agriculture, Forest Service, Southern Forest Experiment Station. doi.org/10.1371/journal.pone.0102920.

Spies, TA, & Franklin, JF. (1991). The structure of natural young, mature, and old-growth Douglas-fir forests in Oregon and Washington. *Wildlife and vegetation of unmanaged Douglas-fir forests*, 91–109.

Szabo, JK, Khwaja, N, Garnett, ST, & Butchart, SH. (2012). Global patterns and drivers of avian extinctions at the species and subspecies level. *PloS one, 7*(10), e47080. doi.org/10.1371/journal.pone.0047080.

Temple, SA, Mossman, MJ, & Ambuel, B. (1979). The ecology and management of avian communities in mixed hardwood-coniferous forests. In RM DeGraff and KE Evans (compilers) *Proceedings of the workshop: Management of north central and northeastern forests for nongame birds* (US Forest Service General Technical Report NC-51). Minneapolis, Minnesota: North Central Forest Experiment Station, US Department of Agriculture, Forest Service. Retrieved from ncrs.fs.fed.us/pubs/gtr/other/gtr_nc051/gtr_nc051l.pdf.

Van Den Meersschaut, D, & Vandekerkhove, K. (2000). Development of a stand-scale forest biodiversity index based on the State Forest Inventory. In: M Hansen, T Burk (eds.) *Integrated tools for natural resources inventories in the 21st century* (General Technical Report NC-212) (pp. 340–50). St. Paul, Minnesota: North Central Forest Experiment Station, US Department of Agriculture, Forest Service.

Watson, J, Freudenberger, D, & Paull, D. (2001). An assessment of the focal-species approach for conserving birds in variegated landscapes in southeastern Australia. *Conservation Biology, 15*(5), 1364–73. doi.org/10.1046/j.1523-1739.2001.00166.x.

The problem of induction in cosmology

ALEX LOMBARD

Abstract

This article problematises a crucial assumption in the methodology of modern cosmology, namely the cosmological principle (CP). Given the physical obstacles and limits peculiar to cosmological observation, cosmological inferences ordinarily take the regularity readily observed in our region of space-time as a basis for the purported self-same regularity of physical processes characteristic of unobservable space-time regions. Just this projection of the locally familiar to the globally unfamiliar is of the essence of the CP. This inductive procedure, however, is beset by serious logical difficulties, an analysis of which is undertaken herein, and the strength of which, it is argued, may constitute grounds for tempering our confidence in cosmological inferences on the nature of the large-scale structure of space-time. The article, all the while, seeks to share in and communicate the excitement at the contributions and prospects of a distinctively philosophical engagement with the conceptual problems of mathematical and theoretical physics, an engagement that is at the very heart of contemporary philosophy of science.

Introduction

Inductive inference has, in various ways, played – and indeed continues to play – a methodologically central role in modern physical cosmology. This type of inference takes past experience as a justification for knowledge of that which has not been experienced. I focus on this aspect of cosmological method, for it is here that the justification of inferences in cosmology faces distinct obstacles over and above the inductive problems that cosmology inherits from the terrestrial (or local) physics upon which it has built.[1] These obstacles chiefly revolve around the legitimacy of the bold extrapolations of experimental observation constitutive of the cosmological principle (CP) and, particularly in the years 1948–1965, of the perfect cosmological principle (PCP), the relevance of which is of more than passing historical interest. This paper argues that in the absence of confidence in the inductive propriety of the CP and

[1] I shall not, therefore, discuss these more classical problems of induction herein.

PCP, the cosmological enterprise is hardly able to get off the ground. The paper proceeds in two stages: first, the CP is defined and two crucial philosophical pitfalls in this definition are identified; second, the CP's problematic temporality is revealed via a discussion of the PCP, and the danger an unsatisfactory treatment of temporality poses to a proper elucidation of the physical dynamics of the universe is emphasised.

The CP and relativistic cosmology

Although only explicitly formulated in 1935 by Milne,[2] the CP is given its first implicit application by Einstein in his 1917 paper applying General Relativity (GR) to the large-scale structure of space-time, the paper which set the foundations for relativistic cosmology.[3] Here Einstein introduces his cosmological constant (λ) which attributes, in the form of an assumption, a structural uniformity – of the distribution of matter and radiation – to the universe necessary for making the latter tractable for GR.[4] Within a decade, Friedmann and Lemaître had independently presented alternative solutions to Einstein's field equations which challenged Einstein's static conception of the universe, providing in effect the mathematical basis for the Big Bang model of the universe.

The subsequent development of Friedmann-Lemaître-Robertson-Walker (FLRW) models in the 1930s explicitly formalised the CP as postulating that the large-scale structure of the universe is (1) spatially homogeneous and (2) spatially isotropic around each space-time point, on *sufficiently large* scales. (1) is defined such that observations obtained from any point in space-time are (approximately) the same as those obtained from any other such point.[5] (2) is defined such that differently directed observations from the same point in space-time are (approximately) the same.[6] Logically, (2) implies (1), but not

[2] Jeremy Butterfield, 'On Under-Determination in Cosmology', *Studies in History and Philosophy of Modern Physics*, vol. 46, May 2014, p. 60 and George Gale, 'Cosmology: Methodological Debates in the 1930s and 1940s', *Stanford Encyclopedia of Philosophy*, 21 June, 2017, sec. 3.2, online.

[3] Claus Beisbart, 'Can We Justifiably Assume the Cosmological Principle in Order to Break Model Underdetermination in Cosmology?' *Journal for General Philosophy of Science*, vol. 40, no. 2, 2009, p. 177.

[4] Crucially, Einstein's cosmological constant was also motivated by his need to correct for what he believed to be an otherwise unacceptable catastrophic implosion of the universe as a result of extant gravitational imbalances distributed throughout the cosmos. See Simon Singh's account in his *Big Bang: The Origin of the Universe*, HarperCollins, London, 2005, ch. 2.

[5] See, for example, Claus Beisbart and Tobias Jung, 'Privileged, Typical, or Not Even That? Our Place in the World According to the Copernican and the Cosmological Principles', *Journal for General Philosophy of Science*, vol. 37, no. 2, 2006, p. 234.

[6] Ibid, p. 242.

vice versa.[7] FLRW models therefore describe an expanding and – because homogeneous – unbounded[8] four-dimensional space-time emerging from a unique Big Bang singularity. Unpacking the CP's definition still further, two philosophical issues emerge: a methodological stress upon observation and a logical question concerning the vagueness in the meaning of the adjective phrase 'sufficiently large'. Insofar as the latter is consequent upon the assumption of the former, I examine each successively.

Unlike in typical terrestrial-focused natural sciences, the CP extrapolates from the observed[9] to regions of the universe long believed to be – ignoring for now the prospects of gravitational wave astronomy[10] – *unobservable* in principle. The threshold of the observable universe is taken to depend upon the propagation of light, the surest – because constant – physical signal and transmitter of cosmic information. The propagation of light, however, is taken to have begun with decoupling, about 380,000 years after the Big Bang. Furthermore, because the expansion velocity of the universe outstrips that of light, it follows that much in the way of light signals is precluded from possibly ever reaching us. The state of the early universe, then, and of the universe outside our past light-cone is logically unobservable.[11] Moreover, the ratio of the size of the observable universe to that of the entire universe is continually decreasing given the latter's expansion velocity. 'To put the point very simply,' writes Butterfield, 'in terms of enumerative induction over spacetime regions: the observable universe is such a small fraction of such regions, that it is risky to claim it is a fair sample.'[12] Perhaps the sample can make up for this quantitative (statistical) deficiency by its fine-grained qualitative detail.

Such a resolution of the sample size problem, however, quickly falters in practice, for the selection of appropriate scales for cosmic observations has been governed not by fine-graining, but by coarse-graining.[13] To conserve the similarity of observations demanded by the CP, spatial homogeneity and spatial isotropy have over the century of modern physical cosmology been measured at larger and larger

---

[7] Butterfield, p. 60; H. Bondi, *Cosmology*, Second edition, Cambridge University Press, Cambridge, 1960, p. 14; and Beisbart and Jung, p. 240.

[8] Beisbart and Jung, p. 233, write: 'It follows immediately from the definition of homogeneity that homogeneous systems cannot have boundaries.'

[9] By today's standards, for instance, FLRW models apply GR 'at length scales 14 orders of magnitude larger than those at which it has been tested' (Quoted in Christopher Smeenk and George Ellis, 'Philosophy of Cosmology', *Stanford Encyclopedia of Philosophy*, 21 December, 2017, sec. 1, online).

[10] See, for example, Joseph Silk and Jens Chluba, 'Next Steps for Cosmology', *Science*, vol. 344, 9 May 2014, pp. 586–7 and BF Schutz, 'Gravitational Wave Astronomy: Delivering on the Promises', arXiv, 17 April 2018, esp. sec. 8. The former talks of 'opening up a new window for exploration of the primeval plasma from which all structure originated', p. 587.

[11] Butterfield, pp. 62–3.

[12] Ibid, p. 63.

[13] Ibid, p. 61 and Beisbart and Jung, p. 242.

scales to blur out the obvious inhomogeneities we observe on smaller scales.[14] Most notably, this scale shift has been forced upon the CP by the discovery first of galaxy clusters, and then of galaxy superclusters;[15] the future discovery of further inhomogeneities requiring yet further upwards adjustment in the CP scales is, logically and physically, all too possible.[16] Today, the scale stands at over 300 million light years.[17] The qualification in the CP's definition that measurements of observations be made on 'sufficiently large scales'[18] is troubling for a scientific principle with as much significance as the CP, for it heaps a vague adverb ('sufficiently') upon a vague scaling adjective ('large'). The indeterminacy in the semantics of 'large' endows the CP with the flexibility to ignore the smaller-scale inhomogeneities. Yet this comes at a high cost, as the following section attests.

The PCP and steady-state cosmology

Although not a sufficient condition of the CP, spatial homogeneity is a necessary condition for it to hold. Measurement scales too large to capture physically significant inhomogeneities therefore risk jeopardising the inductive legitimacy of the cosmological enterprise. At issue here is the *temporality* of physics, which in 1948 Bondi and Gold made a central dissatisfaction with the CP and, by extension, of the Big Bang model of the universe.[19] To guarantee the universality of the laws of physics, they argued, the structure of the universe itself – 'which depends upon the physical laws'[20] – needed to be rid of those temporal distinctions by which any atemporal physics seeking to describe it would have to be modified. But the cardinal principle of science, they tell us, is the repeatability of experiments in the formulation of the laws of nature, which implies the *irrelevance* of the place *and* time of experimentation.[21]

The CP's commitment to spatial isotropy, Bondi and Gold believed, rightly does away with spatial contingency but does not do likewise concerning temporal contingency. They therefore proposed an

---

[14] Butterfield, p. 61.

[15] Ibid, p. 63.

[16] This is especially so given the observational constraints we face in studying the cosmos and, relatedly, the limited range of observational techniques at our disposal in cosmology. The progress of the latter shall hopefully reduce the former and give us a better 'view' of the universe. The advent of gravitational wave astronomy, to use a very contemporary example, may allow us to observe more deeply into the cosmos, and bring with it a new picture of the (in)homogeneity of the large-scale structure of space-time.

[17] Butterfield, p. 63 and Beisbart, p. 199.

[18] Quoted in Butterfield, p. 60.

[19] H. Bondi and T. Gold, 'The Steady-State Theory of the Expanding Universe', *Monthly Notices of the Royal Astronomical Society*, vol. 108, no 3, 1948.

[20] Ibid, p. 254.

[21] Bondi and Gold, p. 252 and Bondi, p. 12.

extension of the CP – called the perfect cosmological principle (PCP) – that fulfils this scientific requirement. The PCP defines the large-scale structure of the universe to be (1) spatially homogeneous and (2) spatially isotropic around each space-time point, on sufficiently large scales, *at all times*.[22] The PCP thus founds a new conception of the universe, namely the Steady-State model: like the Big Bang model, it is spatially unbounded, because of the homogeneity condition of the PCP; unlike the Big Bang model, however, it is also temporally unbounded, because of the 'at all times' stipulation of the PCP.

The Steady-State model posits an eternal space-time manifold, not one that emerged from a Big Bang singularity. By 1948, however, breakthroughs in the understanding of nucleosynthesis had been made by Hoyle, soon to become the most notable proponent of the Steady-State theory.[23] Together with Bondi and Gold, Hoyle demonstrated that a Steady-State universe could save the appearances of an expanding universe – the overwhelming observational evidence of which stretches back to Hubble's 1929 discovery of the red-shift in the spectral lines of receding galaxies – without conceding the universe's catastrophic emergence from a Big Bang singularity. This they did by positing a speculative[24] physical process of 'continual creation of matter'[25] via nucleosynthesis, which ensured that within each PCP scale, the same distribution of matter and radiation would be preserved *over time*.[26] Until 1965, the Big Bang and Steady-State models were each capable of fitting the observational evidence, but the latter's theoretical advantage remained its immunity to the logical problem of projecting – via inductive inference – a temporally-contingent physics (the physics developed over the past 400 years) onto a temporally-evolving universe.

With the discovery of the cosmic microwave background radiation in 1965, the Steady-State model – which predicted its *inexistence* – was taken to be empirically refuted, and the Big Bang model – which

---

[22] Bondi and Gold, pp. 254–5.

[23] See ch. 4 of Singh's account.

[24] Speculative, because it openly violated the principle of conservation of energy. But Bondi and Gold defended the physical hypothesis thus:

> In interpreting the universe as stationary we have to assume that such a process of creation is operative; we have to infringe the principle of hydrodynamic continuity. But this principle is not capable of experimental verification to such a precision, and this infringement does not constitute a contradiction with observational evidence. It is true that hydrodynamic continuity has been regarded as an unqualified truth and not as an approximation to physical laws, but this was merely a bold simplifying extrapolation from evidence. Hydrodynamic continuity is no doubt approximately true but this does not compel us to assume that it holds without any deviation whatever. In the conflict with another principle which is much more far-reaching and capable of making many more statements about the nature of the universe and the applicability of physical laws, there is no reason for upholding the principle of continuity to an indefinite accuracy, far beyond experimental evidence.' (Quoted in Bondi and Gold, p. 256.)

[25] Quoted in Bondi, p. 143.

[26] Ibid, p. 143 and Singh, ch. 4. The expansion of the universe, otherwise, would introduce a temporal diffusion of radiation and matter across each PCP scale.

predicted its existence – was considered validated to an unprecedented extent. Indeed, it very soon thereafter came to be referred to as the Standard Model (SM) in cosmology.[27] The SM's vindication, however, resulted not from a resolution of the temporality problem, but rather to its impressive empirical corroboration. Logically, the SM's commitment to the universality of the laws of physics formulated over the past 400 years sits uncomfortably with its description of a temporally inhomogeneous universe. The SM's admission that there is little warrant for applying the supposedly universal laws of physics to the very early universe[28] (and effectively none for the first 10^{-11} s after the Big Bang) is an explicit reminder that the spectre of temporality continues to haunt the SM.[29] And yet, what is at stake here is the very nature of the physical dynamics of the universe. The concern is that matter inhomogeneities – say at the time of decoupling – filtered out of a coarse-grained CP scale might be crucial input data in the elaboration of a relevant micro-dynamics. This comprises the physical mechanisms informing the evolution of these inhomogeneities into the various galaxies, galaxy clusters and galaxy superclusters that we observe today.[30] Insofar as the latter are governed by a macro-dynamics – namely, GR as presented in FLRW models[31] – the risk here is of the micro-dynamics not linking up, commuting, with the macro-dynamics: the 'question is whether this macro-dynamics is induced by a reasonable coarse-graining procedure on the unknown, myriadly complex, micro-dynamics'.[32] A half-century of work on the SM, in sum, has brought it impressive observational corroboration,[33] but little in the way of progress on the (inductive) logical front.

---

[27] Coined by Weinberg in 1972. (Chris Smeenk, 'Philosophy of Cosmology', in Robert Batterman, ed., *The Oxford Handbook of Philosophy of Physics*, Oxford University Press, Oxford, 2013, p. 609).

[28] Namely, the first hundredth of a second after the Big Bang.

[29] Butterfield, p. 57n.

[30] Ibid, p. 65.

[31] Ibid, p. 65.

[32] Ibid, p. 65. A simple counterfactual analogy may illustrate the problem. Suppose that an extra-terrestrial biologist were to land at a nursing home on Earth and encounter for the first time a few hundred adult humans older than 50, having no previous knowledge of human biology. The biologist quickly notices and studies the phenomenon of sarcopenia, or age-related muscle degeneration. The biologist observes all sample subjects undergoing the same types and rates of muscle degeneration, and subsequently infers that sarcopenia is an intrinsic feature of human biology *at all stages of life*, blithely unaware that the mechanism in question only begins beyond the age of 50. Here, the sample size was too coarse-grained to allow relevant inhomogeneities to inform sound inferences about the mechanisms of human biology. The biologist is tripped up by the absence of a *temporally sensitive* biology. The absence is costly: the biologist's inferences fail to capture the fact that *precisely the opposite* of what they conclude occurs at other stages of life.

[33] See, for instance, Singh's account of the decades-spanning COBE mission, in his ch. 5.

Coda

The problem of induction in cosmology cuts across the realism/instrumentalism debate in the philosophy of science. Realism, whereby scientific – cosmological, in this case – theories are, roughly, taken to be representing the physical world, is obviously affected by the problems identified above. Less obvious is the fact that instrumentalism, too, is unable to remedy these problems. For even instrumentalism, whereby scientific theories are taken to be computational devices or instruments for making predictions, presupposes in this case cosmological observations to verify predictions. The interpretation of cosmological observations, however, often employs the very assumptions the observations are expected to corroborate.[34] Most notably, cosmological observation's reliance upon light propagation is problematised by the fact that our physical understanding of the latter is largely derived from *homogeneous* experimental space-time regions,[35] and further by the fact that inhomogeneities are known to distort it.[36] Vicious logical circles thus impinge upon any innocent recourse to instrumentalism to transcend the problem of induction in cosmology.

Bibliography

Beisbart, C, 'Can We Justifiably Assume the Cosmological Principle in Order to Break Model Underdetermination in Cosmology?' *Journal for General Philosophy of Science*, vol. 40, no. 2, 2009, doi.org/10.1007/s10838-009-9098-9

Beisbart, C, and Jung, T, 'Privileged, Typical, or Not Even That? Our Place in the World According to the Copernican and the Cosmological Principles', *Journal for General Philosophy of Science*, vol. 37, no. 2, 2006, doi.org/10.1007/s10838-006-9015-4

Bondi, H, and Gold, T, 'The Steady-State Theory of the Expanding Universe', *Monthly Notices of the Royal Astronomical Society*, vol. 108, no. 3, 1948, doi.org/10.1093/mnras/108.3.252

Bondi, H, *Cosmology*, Second edition, Cambridge University Press, Cambridge, 1960

Butterfield, J, 'On Under-Determination in Cosmology', *Studies in History and Philosophy of Modern Physics*, vol. 46, May 2014

[34] Beisbart and Jung, p. 246 and Beisbart, p. 183. Consider for instance Bondi's admission in the preface to Part 2 – on observational evidence – of his influential textbook on cosmology, which reads, in part: 'In the discussion of the observational evidence the cosmological principle is used. This procedure seems to be justified owing to the wide agreement on the validity of the cosmological principle and also because its absence would make any interpretation almost impossible.' (Quoted in Bondi, p. 17).

[35] Beisbart and Jung, p. 246 and Beisbart, p. 184.

[36] Butterfield, p. 65.

Gale, G, 'Cosmology: Methodological Debates in the 1930s and 1940s', *Stanford Encyclopedia of Philosophy*, 21 June 2017, plato.stanford.edu/entries/cosmology-30s/

Schutz, BF, 'Gravitational Wave Astronomy: Delivering on the Promises', arXiv, 17 April 2018, royalsocietypublishing.org/doi/abs/10.1098/rsta.2017.0279

Silk, J, and Chluba, J, 'Next Steps for Cosmology', *Science*, vol. 344, 9 May 2014, doi.org/10.1126/science.1252724

Singh, S, *Big Bang: The Origin of the Universe*, HarperCollins, London, 2005

Smeenk, C, 'Philosophy of Cosmology', in Robert Batterman, ed., *The Oxford Handbook of Philosophy of Physics*, Oxford University Press, Oxford, 2013

Smeenk, C, and Ellis, G, 'Philosophy of Cosmology', *Stanford Encyclopedia of Philosophy*, 21 December 2017, plato.stanford.edu/entries/cosmology/

The role of human factors in airport baggage screening

CHENGHAO YU

Abstract

We all spend much time visually scanning our surrounding environment. Visual searching is particularly crucial for public safety when it comes to those occasions that can have fatal consequences. Human factors that influence decision-making in target detection are, however, widely considered as determinants for whether there are prohibited items or not. Cognitive performance of security screening can be influenced by the experiences of the searcher, but also can be affected by the task itself, which is closely associated with principles of perception, visual attention and working memory. This paper critically examines five aspects of the searching task: low target salience; the unknown target set; simultaneous search for all types of targets; the possible occurrence of multiple targets; and low target prevalence. It highlights the challenges of baggage screening and proposes two means for improving the accuracy and the efficiency of the task: operational strategies and off-the-job training. Further studies are needed to confirm whether those strategies could strike a balance between efficacy and efficiency for airport baggage screening in a real-life setting.

Introduction

Whether looking for a mobile phone at home, a file in the folder, or a friend in the crowd, 'much of our life is spent searching for information relevant to the task in hand' (Peterson, Kramer, Wang, Irwin, & McCarley, 2001). When search tasks have life-or-death implications, it is imperative to underscore that the screener performs a search with greater accuracy and efficiency. Baggage screenings at airports constitute a critical search task. While the efficiency of baggage scanning has been significantly improved in both accuracy and speed by technological advances (e.g., X-ray scanning), the ultimate outcome (e.g. the missing rate of the target) is largely attributed to human factors (Biggs & Mitroff, 2014). Paradoxically, the screener is either the strongest or weakest link for most of the screening tasks (Schwaninger, 2006). Currently, it is still not feasible to have such an advanced artificial detection system that can detect the illicit items as quickly and reliably as humans do (Schwaninger, 2005, 2006).

A baggage screener's task is to identify the prohibited items (the target) on an X-ray display that contains a number of other items (the distractors), a procedure requiring attention and working memory (Goldstein, 2010). As the task itself greatly influences cognitive performance in baggage scanning, each

component of the task will be discussed in turn: 1) low target salience; 2) the unknown target set; 3) simultaneously searching for all types of targets; 4) the possible occurrence of multiple targets; and 5) low target prevalence. In doing so, the paper uses a cognitive-psychological approach to highlight the challenges that these task components present for baggage screeners. It offers two types of procedural solutions for improving object recognition and visual attention that do not excessively undermine the efficiency of baggage screening: operational strategies and off-the-job training.

Target salience

Target salience can be enhanced by employing visual cues and camouflaged targets training, and having security staff physically rotate bags prior to screening. Low visibility of the banned items in airport security screening is attributed to three challenging target-based factors: physical distinctiveness; physical orientation; and target–nontarget similarity (Biggs, Adamo, & Mitroff, 2014; Biggs & Mitroff, 2014; Schwaninger, 2005). First, the visibility of the target has been heavily determined by its physical distinctiveness (e.g., shape and size) in a display (Schwaninger, Michel, & Bolfing, 2007). The material proprieties of a target largely interact with the physical identity, which could make it stand out from the screening image. This is attributed to X-ray screening machines mainly differentiating material types by pseudocolouring, such as light orange indicating non-metallic material (see Figure 1).

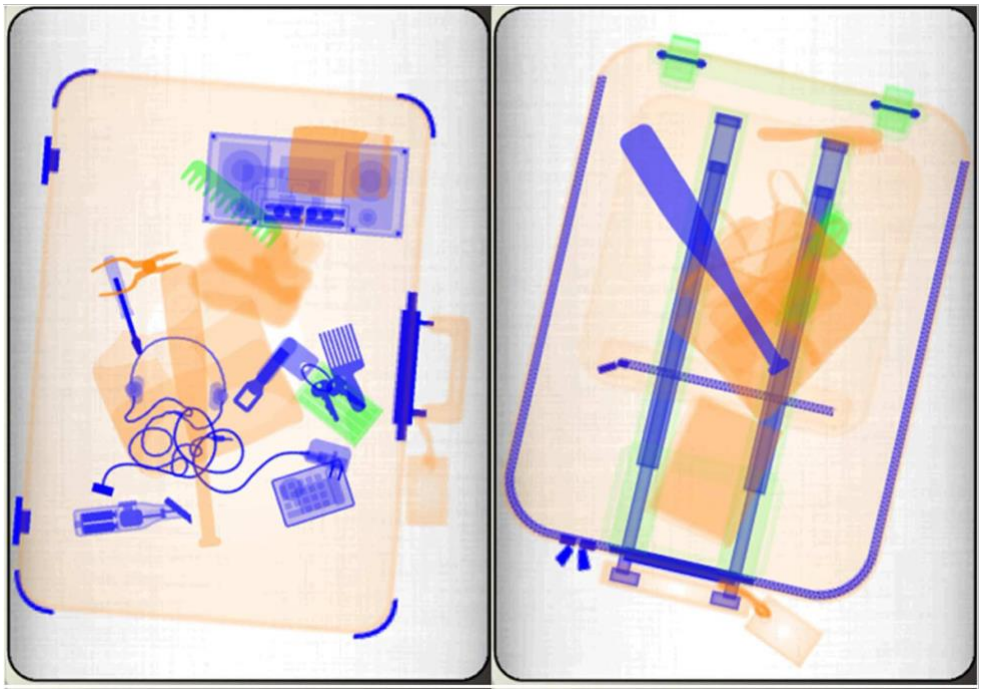

Figure 1: Examples of two different baseball bats: one with lower target visibility (left, wooden baseball bat, coloured light orange) and one with higher target visibility (right, metallic baseball bat, coloured blue)

Source: Biggs et al., 2014.

To mitigate the problem with physical distinctiveness when looking at the baggage image, an operational adjustment and a routine exercise are proposed. That is, setting up a target template provided with detailed visual rather than semantic information, and separating targets from the background via camouflaged target training (Chen & Hegdé, 2012; Vickery, King, & Jiang, 2005). Instead of giving a list of named prohibited items to the screeners, providing detailed visual cues may be more informative. For example, Vickery et al.'s (2005) study not only stressed the importance of prior knowledge of the target in a visual search but also highlighted that the reaction time of a visual cue will be faster than a semantic cue when accuracy is comparable. Matching the target with a specific template (mental representation), however, could compromise the efficiency of visual search (Bravo & Farid, 2014).

The target–background similarity could also profoundly influence a target's physical distinctiveness. In many real-world search occasions such like airport baggage screening, targets and distractors are difficult to segment from the background, which becomes an efficiency cost during the visual search (Boot, Neider, & Kramer, 2009). In this case, an X-ray image could contain items that the machine accidentally filters into the background by pseudocolouring (see Figure 1, where the target of the wooden baseball bat has been faded into the background because of its non-metallic material). This is a sound example of how searching for camouflaged targets involves imperfect segmentation between targets and background (Neider, Boot, & Kramer, 2010). There are nonetheless promising studies to improve the training of camouflaged searches (Boot, Neider, & Kramer, 2009; Neider et al., 2010). In one study, camouflaged-trained participants not only demonstrated improved search performance with fast response times in target recognition, but also fewer eye movements when searching novel camouflaged targets (Boot, Neider, & Kramer, 2009).

Considering the viewpoint-dependent theory proposing that the accuracy and speed of object recognition are affected by the changes in viewpoint, the visibility of illegal items could be influenced by the physical orientation (Tarr, 1995; Tarr & Bülthoff, 1995). This often happens when a banned item is intentionally placed in a particular orientation so that screeners are difficult to detect from their points of view (Bolfing, Halbherr, & Schwaninger, 2008). Having security personnel manually rotate bags into certain positions prior to screening, however, could be a viable solution (Koller, Hardmeier, Michel, & Schwaninger, 2008).

The final factor to consider in target salience is the target–distractor similarity. Research on target distractors, such as feature integration theory and subsequent development, highlights how visual task performance is impaired as the similarity between targets and distractors increases and the similarity among non-targets decreases (Duncan & Humphreys, 1989; Treisman & Sato, 1990). The prohibited items are likely to be missed when surrounded by other physically or conceptually similar but permissible items (Duncan & Humphreys, 1989; Proulx & Egeth, 2006). It is imperative to stress the pitfalls of conceptual similarity when looking at two physically similar items during a routine briefing. Illicit drugs would be less detectable when blending in with legally prescribed drugs, or a bottle of water

versus a bottle of gasoline (Neider, Boot, & Kramer, 2010). This similarity effect might be modulated by using a decision-tree tool accompanied by a visual template to avoid the conceptual traps, as well as having screeners to categorise the non-targets under the same category during a large set-size task. The downside of these procedures is the prolonged searching time as priming effects do occur when targets are absent (Wilschut, Theeuwes, & Olivers, 2014).

Unknown target set

To deal with the problem of the unknown target set, off-job training with a diverse category as well as consistent searching strategies appear to be promising. The unknown target set is twofold in the context of airport baggage screening. First, it accounts for security screeners having no prior and specific knowledge about what items they will be looking at during any given scanning task (Vickery, King, & Jiang, 2005). Second, the combination of items during each screening is unknown, such as the features of combination and set sizes (Bolfing et al., 2008). Accordingly, detection training with diverse categories is introduced to ease the former difficulty , since training with variability can improve screeners' abilities to identify novel objects (Gonzalez & Madhavan, 2011). Gonzalez and Madhavan (2011) indicated that training with high diversity of categories (e.g., knives, guns, scissors, glass objects and metal tools) produced: a higher hit rate; a lower rate of false alarms; and a faster detection time during later detection of novel dangerous targets compared with those who were trained with low diversity of categories (e.g., knives only).

Regarding the unknown combination, consistent visual search strategies could be adopted to offset the memory burden imposed by the unknown target set. That is, a short-term memory burden of recalling what area of the image has or has not been attended occurs when a screener randomly scans the display. Biggs, Cain, Clark, Darling, and Mitroff (2013) observed a higher accuracy of target detection in the professional searchers who have consistent searching behaviour (e.g., scanning from left to right in a row with each display), as compared to those who are less consistent. The study underlines the effectiveness of consistent top-down control to visual search improvement. This strategy enables searchers to reallocate their limited cognitive resources from short-term memory to object recognition.

Simultaneously searching for all types of targets

Given that the current practice in X-ray baggage scanning is to implement a screener to search for all types of targets (e.g., knives, guns, drugs and metal tools) simultaneously in each display, both divided attention and selective attention are under strain for possibly missing any subset (Goldstein, 2010). Divided attention is involved in attending simultaneously to two or more spatially separate locations (Goldstein, 2010). Selective attention refers to selectively processing relevant sensory information for enhanced processing while filtering out the irrelevant visual noise (Underwood & Everatt, 1996). Guided baggage searches are more likely to generate certain inference between top-down requirements

for each type of target and result in inappropriate target representations (Wolfe, Cave, & Franzel, 1989). The cognitive burden in searching for more than one target category could, however, be reduced via dividing a search among multiple searchers across target types (Menneer, Barrett, Phillips, Donnelly, & Cave, 2007). Instead of having one screener to look for both knives and explosives, it could be more efficacious to have one screener to only search for knives and another screener only look for explosives. Whether or not there is enough staff resources to divide tasks should be considered beforehand.

Possible occurrence of multiple targets

Adopting time flexibility and a rerun search by a different screener can be useful for tackling the possible appearance of multiple targets. When multiple targets appear in a single search array, a searcher who finds one target is likely to miss the remaining targets thereafter. This type of error has been referred to as 'subsequent search misses' (SSM) in cognitive psychology literature (Adamo, Cain, & Mitroff, 2013). Recent findings suggest that SSM errors might be attributed to two underlying mechanisms: the perceptual set bias and depletion of cognitive resources (Fleck, Samei, & Mitroff, 2010). The former mechanism assumes that searchers are more susceptible to SSM when subsequent targets are neither physically nor categorically consistent with the found targets. For example, a screener who found a knife is biased to search for additional knives and may possibly miss a dissimilar target – such as an explosive. Meanwhile, the cognitive resources (e.g., visual attention and working memory) are draining due to constantly tracking the identities and locations of found targets (Adamo et al., 2013; Cain & Mitroff, 2013).

Three operational improvements are recommended for reducing SSM in the likelihood of multiple targets. In Adamo, Cain, and Mitroff's 2015 study, more of these errors were observed while looking for additional targets under time pressure. Therefore, when possible, the adjustment of the strict time limit of the screening task is anticipated to alleviate the adverse effects of SSM. Another simple procedure would be to conduct a second search right after the suspicious items have been found and removed, and having a different screener perform the rescan would further improve detection rate (Cain, Biggs, Darling, & Mitroff, 2014). The efficacy of the rerun is achieved by dividing a multiple-target search into several single-target searches, as the memory burden of found targets can be eased (Cain & Mitroff, 2013). Additionally, a little 'pause' can be added during the screening task because searching a moving display can result in increased SSM errors, as compared to a static image (Stothart, Clement, & Brockmole, 2017).

Low target prevalence

Low-prevalence target effects can be improved by providing a short burst of higher prevalence search with feedback; monetary incentives; and knowledge about misses through false feedback. The finding that rare targets are more likely to be missed, also known as the low-prevalence effect, has received

increasing attention from cognitive researchers (e.g., Fleck & Mitroff, 2007; Van Wert, Horowitz, & Wolfe, 2009; Wolfe, Horowitz, & Kenner, 2005). This means that miss rates are much higher when the presence of targets is very infrequent (Fleck & Mitroff, 2007). Compared to the miss rate of 0.09 at the high prevalence of 50 per cent, the miss rate jumped to 0.32 at the low prevalence of 2 per cent (Van Wert, Horowitz, & Wolfe, 2009). This perceptual effect is very robust and persistent in our everyday visual search, despite being subjected to the criterion adjustment of detection (Fleck & Mitroff, 2007).

Essentially, the criterion of baggage screening aims to maximise the likelihood of a 'hit' in searching prohibited items and to minimise the chance of 'miss' (Lynn & Barrett, 2014). This means adopting a liberal criterion for reporting 'yes', even though this sometimes results in a false alarm, and a more conservative criterion for 'no' responses (Lynn & Barrett, 2014). Wolfe et al.'s 2007 study suggested that 'absent' responses of targets are slower than 'present' responses in a high-prevalence search; however, 'absent' responses are faster than 'present' responses in low-prevalence conditions. Contradictory results were presented on whether the effect could be moderated by giving searchers the option to correct their previous response, such as rerunning the bags (e.g., Fleck & Mitroff, 2007 vs Van Wert et al., 2009).

Regardless of such contradictions, there are three alternatives that may help to promote the detection rate in rare target search. First, Wolfe et al. (2007) reported that providing a short burst of higher prevalence search with feedback can encourage searchers to maintain a high-prevalence criterion during a low-prevalence search. For instance, a high-prevalence of the training set with feedback can be presented during the middle of a screener's working shift (Wolfe et al., 2007). This study also highlighted the important effects of change in the monetary incentives on the searchers' criterion for low-prevalence searching (e.g., commissions are associated with the number of prohibited items found) (Wolfe et al., 2007). Further, there is evidence to suggest that more infrequent targets could be detected by increasing the searchers' perceived misses via false feedback (Schwark, Sandry, MacDonald, & Dolgov, 2012).

Conclusion

This paper has employed a cognitive-psychological framework to assess the five significant challenges imposed by airport baggage screening tasks. The difficulty with target salience and the low-prevalence effects constitute the two greatest challenges, due to their inherent complexity compared to the other three. Strategies such as a specific visual template, camouflaged targets training and categorising non-targets could, however, be promising, even though sometimes inefficient. To deal with unknown target set, detection training with diverse categories and consistent visual search strategies are most likely to strike a balance between efficacy and efficiency for baggage screening. The problem with simultaneously searching for all types of targets could be moderated by dividing a search among

multiple searchers across target types. Whether there are enough staff resources to do this, though, should be considered.

In order to address the possible appearance of multiple targets, it is advised that airport security adopt a flexible time limit of the screening task and conduct a rerun search by a different screener. The problem of low-prevalence targets could be resolved by employing a short burst of higher prevalence search with feedback. Subsequently, a change in the payoff matrix and heightening the searchers' sensitivity to misses via false feedback could be two other feasible options. Of course, the practicality of these proposed solutions might be constrained as they are mostly derived from laboratory-based studies. Further studies should, therefore, focus on the effectiveness and efficiency of these procedural solutions and training in real-life baggage screening scenarios at the airport. Again, technology may facilitate visual searching and object recognition in general, but when it comes to those inherently complex as well as socially important search tasks, human factors still are the strongest or weakest link in security scanning.

References

Adamo, SH, Cain, MS, & Mitroff, SR. (2013). Self-induced attentional blink: A cause of errors in multiple-target search. *Psychological Science, 24*(12), 2569–74. doi:10.1177/0956797613497970.

Adamo, SH, Cain, MS, & Mitroff, SR. (2015). Satisfaction at last: Evidence for the 'satisfaction' hypothesis for multiple-target search errors. *Visual Cognition, 23*(7), 821. doi:10.1080/13506285.2015.1093248.

Biggs, AT, Adamo, SH, & Mitroff, SR. (2014). Rare, but obviously there: Effects of target frequency and salience on visual search accuracy. *Acta psychologica, 152*, 158–65. doi:10.1016/j.actpsy.2014.08.005.

Biggs, AT, Cain, MS, Clark, K, Darling, EF, & Mitroff, SR. (2013). Assessing visual search performance differences between Transportation Security Administration officers and nonprofessional visual searchers. *Visual Cognition, 21*(3), 330–52. doi:10.1080/13506285.2013.790329.

Biggs, AT, & Mitroff, SR. (2014). Improving the efficacy of security screening tasks: A review of visual search challenges and ways to mitigate their adverse effects. *Applied Cognitive Psychology, 29*(1), 142–48. doi:10.1002/acp.3083.

Bolfing, A, Halbherr, T, & Schwaninger, A. (2008). How image based factors and human factors contribute to threat detection performance in X-ray aviation security screening. In A Holzinger (ed), *HCI and usability for education and work. USAB 2008. Lecture notes in computer science*, *vol 5298*. Berlin, Heidelberg: Springer. doi.org/10.1007/978-3-540-89350-9_30.

Boot, WR, Neider, MB, & Kramer, AF. (2009). Training and transfer of training in the search for camouflaged targets. *Attention, Perception, & Psychophysics, 71*(4), 950–63. doi:10.3758/APP.71.4.950.

Bravo, MJ, & Farid, H. (2014). Informative cues can slow search: The cost of matching a specific template. *Attention, Perception, & Psychophysics, 76*(1), 32–39. doi:10.3758/s13414-013-0532-z.

Cain, MS, Biggs, AT, Darling, EF, & Mitroff, SR. (2014). A little bit of history repeating: Splitting up multiple-target visual searches decreases second-target miss errors. *Journal of Experimental Psychology: Applied, 20*(2), 112–25. doi:10.1037/xap0000014.

Cain, MS, & Mitroff, SR. (2013). Memory for found targets interferes with subsequent performance in multiple-target visual search. *Journal of Experimental Psychology, 39*(5), 1398–408. doi:10.1037/a0030726.

Chen, X, & Hegdé, J. (2012). Learning to break camouflage by learning the background. *Psychological Science, 23*(11), 1395–403. doi:10.1177/0956797612445315.

Duncan, J, & Humphreys, GW. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433–58. doi:10.1037/0033-295X.96.3.433.

Fleck, MS, & Mitroff, SR. (2007). Rare targets are rarely missed in the correctable search. *Psychological Science, 18*(11), 943–47. doi:10.1111/j.1467-9280.2007.02006.x.

Fleck, MS, Samei, E, & Mitroff, SR. (2010). Generalized 'satisfaction of search': Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied, 16*(1), 60–71. doi:10.1037/a0018629.

Goldstein, EB. (2010). *Sensation and perception* (8th ed.). United States; Australia; Wadsworth Cengage Learning.

Gonzalez, C, & Madhavan, P. (2011). Diversity during training enhances the detection of novel stimuli. *Journal of Cognitive Psychology, 23*(3), 342–50. doi:10.1080/20445911.2011.507187.

Koller, SM, Hardmeier, D, Michel, S, & Schwaninger, A. (2008). Investigating training, transfer and viewpoint effects resulting from recurrent CBT of X-Ray image interpretation. *Journal of Transportation Security, 1*(2), 81–106. doi:10.1007/s12198-007-0006-4.

Lynn, SK, & Barrett, LF. (2014). 'Utilizing' Signal Detection Theory. *Psychological Science*, 25(9), 1663–73. doi:10.1177/0956797614541991.

Menneer, T, Barrett, DJK, Phillips, L, Donnelly, N, & Cave, K. R. (2007). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology, 21*(7), 915–32. doi:10.1002/acp.1305.

Neider, MB, Boot, WR, & Kramer, AF. (2010). Visual search for real world targets under conditions of high target–background similarity: Exploring training and transfer in younger and older adults. *Acta Psychologica, 134*(1), 29–39. doi:10.1016/j.actpsy.2009.12.001.

Peterson, MS, Kramer, AF, Wang, RF, Irwin, DE, & McCarley, JS. (2001). Visual search has memory. *Psychological Science, 12*(4), 287–92. doi:10.1111/1467-9280.00353.

Proulx, MJ., & Egeth, HE. (2006). Target-nontarget similarity modulates stimulus-driven control in visual search. *Psychonomic Bulletin & Review, 13*(3), 524–29. doi:10.3758/BF03193880.

Schwaninger, A. (2005). Increasing efficiency in airport security screening. *WIT Transactions on the Built Environment, 82*, 407–16. doi:10.2495/SAFE050401.

Schwaninger, A. (2006). *Airport security human factors: From the weakest to the strongest link in airport security screening.* Paper presented at the 4th International Aviation Security Technology Symposium, Washington, DC, USA.

Schwaninger, A, Michel, S, & Bolfing, A. (2007). *A statistical approach for image difficulty estimation in x-ray screening using image measurements.* Paper presented at the the 4th symposium on Applied perception in graphics and visualization, Tubingen, Germany. doi.org/10.1145/1272582.1272606.

Schwark, J, Sandry, J, MacDonald, J, & Dolgov, I. (2012). False feedback increases detection of low-prevalence targets in visual search. *Attention, Perception, & Psychophysics, 74*(8), 1583–89. doi:10.3758/s13414-012-0354-4.

Stothart, C, Clement, A, & Brockmole, JR. (2017). Satisfaction in motion: Subsequent search misses are more likely in moving search displays. *Psychonomic Bulletin & Review, 25*(1), 409–15. doi:10.3758/s13423-017-1300-1.

Tarr, MJ. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2*(1), 55–82. doi:10.3758/BF03214412.

Tarr, MJ, & Bülthoff, HH. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of experimental psychology. Human perception and performance, 21*(6), 1494–505. doi:10.1037/0096-1523.21.6.1494.

Treisman, A, & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 459–78. doi:10.1037/0096-1523.16.3.459.

Underwood, G, & Everatt, J. (1996). Automatic and controlled information processing: The role of attention in the processing of novelty. In O Neumann & AF Sanders (eds), *Handbook of perception and action, vol. 3: Attention*. (pp. 185–227). San Diego, CA, US: Academic Press.

Van Wert, MJ, Horowitz, TS, & Wolfe, JM. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception, & Psychophysics, 71*(3), 541–53. doi:10.3758/APP.71.3.541.

Vickery, TJ, King, L-W, & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision, 5*(1), 8–8. doi:10.1167/5.1.8.

Wilschut, A, Theeuwes, J, & Olivers, CNL. (2014). Priming and the guidance by visual and categorical templates in visual search. *Frontiers in Psychology, 5*, 148. doi:10.3389/fpsyg.2014.00148.

Wolfe, JM, Cave, KR, & Franzel, SL. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 419–33. doi:10.1037/0096-1523.15.3.419.

Wolfe, JM, Horowitz, TS, & Kenner, NM. (2005). Rare items often missed in visual searches. *Nature, 435*(7041), 439. doi.org/10.1038/435439a.

Wolfe, JM, Horowitz, TS, Van Wert, MJ, Kenner, NM, Place, SS, & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General, 136*(4), 623–38. doi:10.1037/0096-3445.136.4.6.

Cross-linguistic influence on Chinese-L2 learners' acquisition of classifiers

JIAHUAN ZHANG

Abstract

This research aims to investigate cross-linguistic influence (CLI) on Chinese-L2 learners' acquisition of classifiers. So far, limited empirical studies have concerned themselves with how L1 affects second language acquisition (SLA), specifically in relation to the acquisition of classifiers. Although many theories and hypotheses have predicted that L1 contributes both positive and negative effects to L2 acquisition, this current research will scrutinise the reasons this is so. I conducted a picture-based composition via dynamic assessment. Two participants with contrasting L1 backgrounds were chosen: Sally and Yui, from New Zealand and Thailand respectively. Findings suggest that L1 both facilitates and hinders SLA, where reasons can be explained through the L1–L2 similarities and differences in classifier system as well as L1 syntactic transfer. These mixed results demonstrate that language acquisition is, predominantly, a particularly complicated process in which other factors such as the linguistic environment, age of acquisition (AoA) and individual differences must be considered within the overall analysis.

Introduction

Researchers have long been paying attention to cross-linguistic influence (CLI) on the procedure of second language acquisition (SLA), which is a key concept under the umbrella term 'language transfer'[1] (Yi, 2013, p. 2372). CLI unfolds how one language can have an effect on another language, contributing both positive (e.g., easiness) and negative (e.g., overproduction, underproduction) influences to learners' L2 acquisition (cf. Anderson, 1983; Flege, 1995; Hyltenstam, 1977; Lado, 1957; Selinker, 1972). Thus far, there is limited validation of CLI on Chinese-L2 speakers' acquisition of classifiers. Typologically, classifiers are a group of morphemes that give categorisation to nouns, being a crucial grammatical category in typical classifier languages, such as Chinese. To illustrate, within a numeral noun phrase (hereafter NP), a classifier must be inserted between the number and the noun in order to suffice Chinese grammaticality. For example, the equivalence of *three tables* in Chinese must be:

[1] 'Language transfer' is also referred to as *L1 interference*, *linguistic inference* and *cross-linguistic influence*. Considering the fact that L2 may also influence L3 acquisition, therefore this article adopts the specified concept *cross-linguistic influence* instead of *language transfer*.

San zhang zhouzi

Three CL. table

'three tables'

This study aims to scrutinise the CLI on classifiers acquisition by adopting dynamic assessment. Data was collected from two Chinese-L2 learners, whose L1s are English and Thai, respectively. Findings suggest that due to L1–L2 similarities and differences, CLI performs both positively and negatively on learners' acquisition of classifiers, through the lens of selecting classifiers and syntactic transfer involving classifiers. This article takes the position that CLI could facilitate or impede L2 learners' acquisition of classifiers. It argues that while due to SLA's complexity *per se*, other potential factors such as linguistic environment, age of acquisition (AoA)[2] and individual differences should also be taken into account in the analysis of language acquisition. I begin with a literature review of Chinese language classifier instruments, before outlining my methodology and findings. I then reflect on the implications for further research on the processes undertaken by Chinese-L2 learners.

Literature review

Cross-linguistic influence

According to Ortega (2014, p. 31), L2 learners have an established L1 capacity before learning L2 and the pre-existing L1 knowledge influences one's L2 acquisition. This phenomenon, generalised as CLI, technically involves a wide range of linguistic content, such as phonology, lexis, syntax, semantics and pragmatics. A string of relevant theories includes Lado (1957)'s contrastive analysis hypothesis (CAH), which suggests that differences and similarities between L1 and L2 predicts difficulty and easiness, respectively, to acquisition. Yet this notion is later disputed by Hyltenstam (1977) and Flege (1995), who argue that similarity may give rise to confusion and therefore results in a negative influence on SLA. Consequently, Anderson (1983) suggests that L1–L2 similarity could have a misleading influence on L2, owing to the principle of 'transferability'[3]. Nevertheless, as early as 1972, Selinker raised the concept of 'interlanguage', denoting that some L1 knowledge is consciously or unconsciously mapped into learner's L2 acquisition, where the trend of unexpected repetitive mappings (i.e., error[4]) is termed as 'fossilisation'. In the discussion section, the current study will apply these notions to provide a case-by-case analysis to the data of participants.

---

[2] Age of acquisition refers to the onset time that learners are immersed in the L2 context, including formal schooling, visiting to L2 country and so on.

[3] *Transferability* means that language knowledge can be transferred into another language in the process of L2 acquisition.

[4] *Error* is an applied linguistic term differing from *mistake*: error refers to a systematic misuse of language, while mistake means an occasional linguistic misuse.

Empirical studies on Chinese classifiers acquisition

At present, a handful of empirical studies have been investigating non-native speakers' acquisition of Chinese classifiers from different perspectives. For instance, Paul and Gruter (2016) probed the approach to learning classifiers from 30 Swedish-L1 speakers; Gao (2010) examined 48 native English speakers' application of classifiers, focusing on the linguistic transfer from L1 prior knowledge; while Kuo (2015) tested 35 Chinese-L2 learners in Taiwan with the aim of exploring the correlation between classifier acquisition and cognitive performance. Despite their different focus, a consistent finding across all of these studies reveals that CLI facilitates but also impedes Chinese-L2 learners' classifier acquisition.

Dynamic assessment

The terminology *dynamic assessment*, coined by Luria (1961), is an initiative achievement of Russian psychologist Lev Vygotsky's research. Dynamic assessment is regarded as a 'pedagogical instantiation of the ZPD' [5] emerging from the dialectical perspective of language learning and teaching (cf. Lantolf, 2009, p. 359; 2013, p. 66). To explain, conducting dynamic assessment is an interactive procedure whereby the examiner guides the examinee individually with scaffolding instructions. By this approach, examinees are able to perform better in the testing as well as learning new knowledge based on their current level. Hence, the final result could be referred to as a learning representation during the assessment procedure.

Methodology

Participants

Data was collected from two Chinese-L2 learners. The first was Sally (anonymised name), aged 21, from New Zealand, who is currently an undergraduate student in Australia. She has already completed a higher-intermediate Chinese language course (lasting two years) as her minor; and her L1, English, is not considered a classifier language. The second participant was Yui (anonymised name), aged 22, a female student from Thailand, who currently studies clinical medicine as an exchange student in China. She studied Chinese for four years in Thailand before studying at a Chinese university; and her L1, Thai, is a typical classifier language.

[5] *Zone of proximal development*: this concept was first purposed by Russian psychologist Lev Vygotsky in 1933, denoting an abstract distance from what a learner can do with assistance to what a learner can do independently in the domain of language acquisition.

Data collection and analysis

Data was collected via a picture-based written task (lasting around 15 minutes) and a 10-minute interview. Participants were asked in advance to write a short picture-based composition.[6] They did not know classifiers were the focus of this task. According to the principle of dynamic assessment, I designed a series of questions[7] before delivering the task. I prepared pens in different colours[8] for participants to mark the trace of testing rounds. Yui's first attempt of the task was almost perfect in applying classifiers, so I did not continue dynamic assessment for her. Instead, I requested her to write another composition in her L1 without referring to the first composition so as to examine whether her acquisition of classifiers is correlated with her L1 knowledge.

Before testing participants, I collected a written sample from a native adult Chinese speaker, as a reference of judgement for the correct use of classifiers. I adopted accuracy and emergence respectively as the benchmarks of acquisition and noticing. Accuracy refers to both syntactic correctness[9] and semantic correctness.[10] Emergence means that participants locate the classifier's position, but they are not able to figure out the correct classifier. In addition, I allowed participants to use *Pinyin*[11] instead of Chinese characters, in case the characters' forms were written wrongly, which would affect the accuracy rate of classifiers.

Findings

Data of Sally

Overall, Sally tended to use 'location+*you*+NP', a structure in which classifiers are normally required. Table 1 presents Sally's dynamic data in sequential rounds, demonstrating how she applied classifiers under my step-by-step instructions.

Table 1: Record of Sally's dynamic assessment

Number	Classifier	Object	1st round	2nd round	3rd round	4th round
1	zhi	dog	√			
2	ge	person	√			
3	ge	cat	×	√		

[6] See Figure 1 in Appendx.

[7] See 'Question formulation in the assessments' in Appendix.

[8] See Figures 2–4 in Appendix: Figure 2, in particular, shows the different colours tracing testing rounds.

[9] i.e., within NP layer, the classifier is placed between the number and the noun.

[10] i.e., the classifier is correctly matched to the noun.

[11] *Pinyin* is the official romanization system for marking pronunciation of standard modern Chinese in mainland China.

4	ge	table	×		√	
5	ge	plant	×		+	
6	ge	fish	×		+	
7	ge	character	–			√

Note: √ indicates the classifier is correct; × is first elicited but is wrong; + is elicited after instructions but is wrong; – is not elicited.

Source: Author's summary of experiment data.

In the first assessment, Sally noticed six classifiers, only one of those is considered to be acquired. In the second assessment, she elicited one misused classifier and revised it into the correct one. In the next round, she figured out three more cases of incorrectness, with two of them still remaining unacquired, and one being correctly amended (see Figure 2 in the Appendix).

In the final assessment, Sally elicited another classifier (not previously noticed) with uncertainty. It seemed to be a coincidence of a correct modification, since she could not locate the classifier by herself but managed to match the appropriate one, *ge* (the generic classifier) for the noun. As a whole, she acquired five classifiers after four rounds' assessments. It is also noted that the data demonstrates a trend of overgeneralisation to the classifier *ge*,[12] because she applied it before each noun when being uncertain.

It is plausible that Sally realised the grammatical importance of classifiers and attempted to apply as many classifiers as she could. Yet, it is clear that she did not recognise the obligatory occurrence of classifier *ge* for 'character' until the final attempt. It could be concluded that Sally has presented a weak (or at least passive) form of classifier acquisition, along with a positive sense of noticing classifiers.

Data of Yui

In contrast to Sally's learning of classifiers, Yui's data indicates a superior mastery of classifiers. Yui finished the task without assistance from dynamic assessment, since her composition is almost error-free. This includes accurate applications of classifiers both in the syntactic layer and the semantic layer.

Table 2: Record of Yui's dynamic assessment

Number	Classifier	Object	1st round
1	ge	person	√
2	zhi	fish	√

[12] *Ge* is a generic classifier, which is applied when the speaker is uncertain which classifier to use. However, it does not mean that people can use *ge* all the time.

3	ge	moment	√
4	zhang	picture	√

Note: √ indicates the classifier is correct.

Source: Author's summary of experiment data.

In total, only four classifiers were elicited in the composition, along with one being misused (i.e., the first classifier *ge*, is technically correct but it sounds awkward to native speakers). I would not, however count this misuse of *ge* as degradation of her classifier acquisition. The reason behind this is that this sentence structure does not require a classifier when the subject is a collective noun. This situation mirrors the issue of conventional usage rather than the acquisition of classifiers.

On the other hand, Yui's syntactic structure is more complicated, which gives rise to avoidance of classifiers. To illustrate, Yui widely adopted the structure of 'Sub+*you*+V+NP'[13] in the composition where classifiers' usage could be discarded without ruining grammaticality *per se*. For example, the following sentence is partially correct because the three classifiers are applied inconsistently before nouns.

Tamen you yang gou, yang mao, hai yang liangzhiyu

They you raise dog, raise cat, and also raise two CL. fish

'They raise dog, cat and two fishes'

Although this sentence is comprehensive for native Chinese speakers, it is difficult to confirm whether Yui thoroughly understands how to apply classifiers in a 'Sub+*you*+V+NP' structure. To explain, it should be noticed that the semantics change slightly in this context[14] though her syntactic structure is acceptable. This can be regarded as evidence of insufficient classifier acquisition through the interface of syntax and semantics. Furthermore, Table 3 illustrates that Yui applied classifiers more frequently in her L1 Thai, with only one classifier (No. 2) appearing in both versions (see Figures 3 and 4 in the Appendix).

Table 3: Yui's application of classifiers in her Thai composition[15]

Number	Noun matching to classifier	Classifier in Chinese version	Classifier in Thai version
1	single person	+	–

[13] Predicate *you* is the focus of this structure. In *you* construction, classifiers are optional in the NP, however, generally native speakers would keep consistent in allocating classifiers: neither repeating classifiers each time nor avoiding all classifiers.

[14] To explain, if classifiers are applied before each noun, this sentence emphasises the different quantities of the animals they have. However, Yui only placed a classifier for 'fish', thereby leading to the interpretation of stressing the quantity of fish but weakening the importance of dog and cat. To some extent this is awkward, considering the setting in the picture, because there is no reason to only emphasise the quantity of fish.

[15] The usage of *ge* here is partially correct.

2	fish	+	+
3	moment	+	–
4	character	+	–
5	cloud	–	+
6	bird	–	+
7	two persons	–	+
8	lives	–	+
9	house	–	+

Note: √ indicates the classifier is correct; × is first elicited but is wrong; + is elicited after instructions but is wrong; – is not elicited.

Source: Author's summary of experiment data.

It is also noted that some classifiers she applied in Thai (Nos. 5–9) have their own counterparts in Chinese. However, she did not apply these classifiers in her Chinese composition. Therefore, I draw a tentative conclusion that Yui created adequate context to demonstrate her knowledge of applying classifiers, while evidence may also point to an avoidance of classifier application.

Discussion

CAH on application of classifiers

By and large, findings from the current study are in line with Lado's (1957) CAH: that L1–L2 similarities lead to easiness but differences predict difficulty. Sally's data is prone to the claim that L1–L2 differences impede her acquisition of classifiers. This is potentially because, compared to Chinese, there are considerably fewer classifiers in English and classifiers are omitted in most contexts. Furthermore, Chinese classifiers are an obligatory component, bounded within numeral NP, and the classifier system is rather elaborate, while classifiers get demoted in English grammar since they are not necessary in numeral NPs.

In Table 4, it is demonstrated that Sally mastered the rule of applying classifiers aptly in the syntactic layer, where there is striking evidence of sufficient noticing of classifiers (from 86% to 100%). It is evident, however, that she has difficulty in mapping the correct classifier for each noun, referring to the final accuracy rate of 71.4%. In the interview, she confirmed that classifier–noun matching is the most challenging part of learning, and her solution is rote memory. Therefore, for Sally it seems to be an ad

hoc scheme to adopt the generic classifier *ge* as an alternative for those classifiers she is unfamiliar with or uncertain about, as demonstrated in Table 1.

Table 4: Rate of accuracy and noticing[16]

Participant	Accuracy	Noticing
Sally	71.4% (28.6%)	100% (86%)
Yui	100% (100%)	100% (100%)

Source: Author's summary of experiment data.

On the other hand, Yui's case is an instance of similarity-easiness in L2 acquisition. In Thai, classifier is a must-be component, as in Chinese, which has to be bounded with the noun it modifies. It is noticed that Yui's pre-knowledge of Thai classifiers facilitates her Chinese classifier acquisition. Through my observation, Yui finished the task without much effort, ending with a perfect accuracy rate and noticing rate of classifiers. Secondly, she also mentioned in the interview that Thai has many classifiers, even though they are encoded in a different cognitive system. Still, being familiar with the usage of Thai classifiers benefits her acquisition of Chinese classifiers, especially in the syntactic layer; and all she needs to do is to memorise the meaning of each Chinese classifier. In a nutshell, similarity of classifiers in Yui's L1 and L2 undeniably simplifies her acquisition of Chinese classifiers.

Similarity–difficulty on syntactic transfer

I perceive that similarities also predict difficulty, consistent with findings from Hyltenstam (1977), Flege (1995) and Anderson (1983). This is reflected in the syntactic layer relevant to classifiers rather than merely phrasal structure. For example, Sally widely adopted a 'location+*you*+NP' structure, which is similar to a 'there be' structure in English. This is a representative case of 'fossilisation' in grammatical pattern as per Selinker (1972). To be more specific, when Sally describes the setting in the picture, she naturally codes sentences with a 'location+*you*+NP' formula, where classifiers are an obligatory element within the NP. To make the sentence correct, she must apply corresponding classifiers. This implicitly requires a higher level of classifier acquisition. To some extent, classifier mis-usage in this context is an implicative transferred result owing to the similarity in L1–L2 syntactic structure. A similar case has been further discussed by Yuan (2015), who observed that L1–L2 similarity of bearing wh-questions hardly provides easiness for English-L1 speakers' Chinese learning. Much to the contrary, it gives rise to difficulty because of the syntactic difference in sequencing wh-questions, which resonates with Anderson (1983)'s notion of negative effect from L1–L2 similarity.

[16] Data in brackets are from the first-round assessment.

Reflection on method

I have given much thought to whether I should elicit the concept of classifier in the question. It appears to be a conundrum. If I asked the participants 'could you please describe this picture with as many classifiers as you can?', then this would probably lead to a confirmation bias in noticing the classifiers. In this case, participants may attempt to use classifiers more frequently than their natural usage, which is very technically directed and it may confound the actual CLI (as it is unnatural). On the other hand, if I did not elicit classifiers (as I did in this project), it is arduous to discern whether or not the participants actually intended to avoid classifiers, and whether or not the avoidance of classifiers merely results from syntactic structure through the data and the interview.

Reflection of analysis

SLA is a complex procedure where ultimate attainment is affected by many variables. These variables correlate with each other and it is difficult to detect or prove which one outperforms the others. Consequently, the analysis of language learning thus far in the current study should consider at least three more factors. As well as emphasising the role of CLI on classifier acquisition, linguistic environment, AoA and individual differences should also be taken into account. Due to word limit, I will briefly mention these three factors as a closing remark.

First, there is the language environment. Sally learns Chinese solely from language classes in Australia, with comparably less exposure to the target language. In contrast, Yui obtained language knowledge in Thailand, where she was immersed in more Chinese culture as well as Chinese language; not to mention the fact that she has also been under considerable language exposure by studying in China. These conditions should have contributed to improve Yui's Chinese language skills. AoA is another critical factor. It is claimed by Birdsong (2005) that AoA is strongly predicative of the end state of SLA; and higher ultimate attainment benefits from younger AoA. Yui started learning Chinese one year earlier than Sally: undeniably her learning length quantitatively exceeds that of Sally. The third factor to consider is individual differences. Although from the interview both participants manifested a strong conation for Chinese learning, it cannot be pinpointed whether cognition factors, such as working memory or aptitude, account for their learning processes and therefore differentiate their ultimate attainment. Considering these implications, further studies regarding Chinese-L2 learners' acquisition of classifiers are suggested to take these factors into account so as to make a more convincing analysis.

Acknowledgements

I would like to thank Dr Susy Macqueen for her expert advice and encouragement throughout this project, as well as Dr Julia Brown for all her tedious work and extensive guidance in revising my drafts. In addition, I appreciate Zirui, who helped me contact the testing participant from Thailand. This pilot project would not have been possible without the support of these lovely people.

References

Andersen, R. (1983). Transfer to Somewhere. *Language Transfer in Language Learning*, *1*, 177–201.

Birdsong, D. (2005). Interpreting Age Effects in Second Language Acquisition. In A Davies & C Elder (Eds), *The Handbook of Applied Linguistics*. (pp. 82–105). Malden, MA: Blackwell.

Flege, JE. (1995). Second Language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Timonium, MD: York Press, 233–77.

Gao, H. (2010). A Study of Swedish Speakers' Learning of Chinese Classifiers. *Nordic Journal of Linguistics, 33*(2), 197–229. doi.org/10.1017/S0332586510000156.

Hyltenstam, K. (1977). Implicational Patterns in Interlanguage Syntax Variation. *Language Learning*, *27*(2), 383–410. doi.org/10.1111/j.1467-1770.1977.tb00129.x.

Kuo, JY. (2015). Cognition-based Multimedia Classifier Learning. *Chinese as a Second Language Research, 4*(1), 23–45. doi:10.1515/caslar-2015-0002. doi.org/10.1515/caslar-2015-0002.

Lado, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Michigan: University of Michigan Press.

Lantolf, JP. (2009). Dynamic Assessment: The Dialectic Integration of Instruction and Assessment. *Language Teaching*, *42*(3), 355–68. doi.org/10.1017/S0261444808005569.

Lantolf. JP. (2013). Sociocultural Theory: a Dialectical Approach to L2 Research. In SM Gass & A Mackey (Eds), *The Routledge Handbook of Second Language Acquisition*. London: Routledge.

Luria, AR. (1961). Study of the Abnormal Child. *American Journal of Orthopsychiatry: A Journal of Human Behavior, 31*, 1–16.

Ortega, L. (2014). *Understanding Second Language Acquisition*. London: Routledge. doi.org/10.4324/9780203777282.

Paul, JZ, & Grüter, T. (2016). Blocking Effects in the Learning of Chinese classifiers. *Language Learning, 66*(4), 972–99. doi:10.1111/lang.12197.

Selinker, L. (1972). Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching, 10*(1-4), 209–32. doi.org/10.1515/iral.1972.10.1-4.209.

Yi, A. (2012). On the Factors Influencing L1 Transfer. *Theory and Practice in Language Studies*, *2*(11), 2372–77. doi.org/10.4304/tpls.2.11.2372-2377.

Yuan, B. (2015). The Effect of Computational Complexity on L1 Transfer: Evidence from L2 Chinese Attitude-bearing Wh-questions. *Lingua*, *167*, 1–18. doi.org/10.1016/j.lingua.2015.09.001.

Appendix

Supplementary

Question formulation in the assessments

1. Please describe what you see in this picture (Figure 1).

Figure 1: Picture of the task.

Source: Author's own design and photograph.

2. Now I will ask you to look over your writing and revise or edit any errors you notice.

3. Can you see an error in this sentence? (When participants are unable to locate error or fill in a classifier in the sentence.)

4. Can you see an error in this noun phrase? (When participants are unable to locate error or fill in a classifier in the noun phrase.)

5. Can you think of what should come before the noun? (When participants are unable to locate error or fill in a classifier between a number and a noun.)

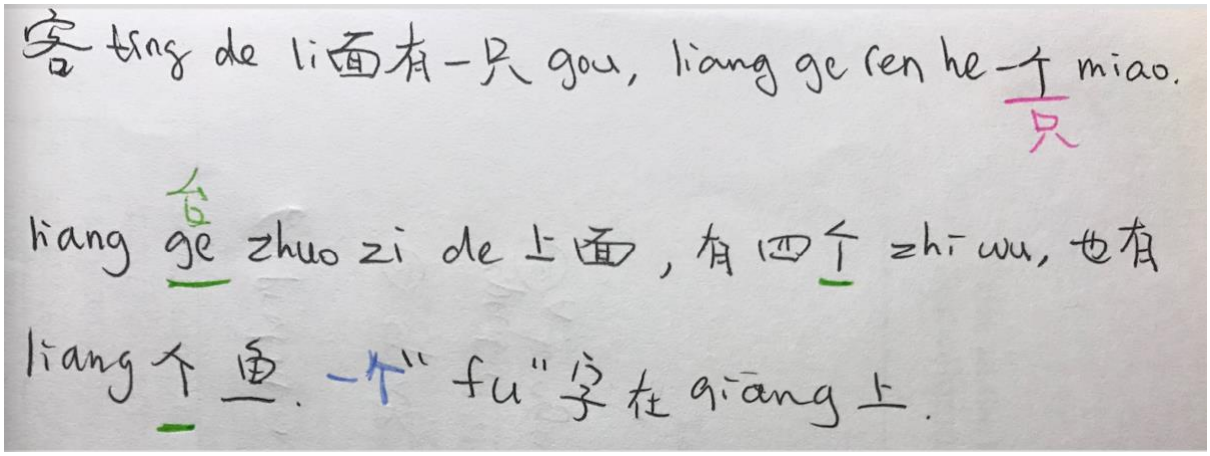

Figure 2: Sally's composition (Chinese).

Source: Author's photograph.

Individual interview with Sally

1.Q: Are classifiers difficult for you? Why?

A: Yes, it is difficult, because in English we seldom use classifiers. And the most difficult part is to memorise classifiers … I know when and where to use classifiers in the sentence, but just have problems in memorising which one is the good one to use …

2.Q: Have you been taught classifiers?

A: Yes, in the courses. They have some part of classifiers teaching sometimes … from very early. in the course you get training like this, you have to talk about them [classifiers] in Chinese, because they are so common …

3.Q: Have you done anything specific in your own self-study to learn classifiers?

A: Emm … normally I treat it like a general study of vocab[ulary]. I often use flash cards as a strategy for me to help me memorise … Not exactly specific method … I learn classifiers mostly by rote memory … to learn the pattern …

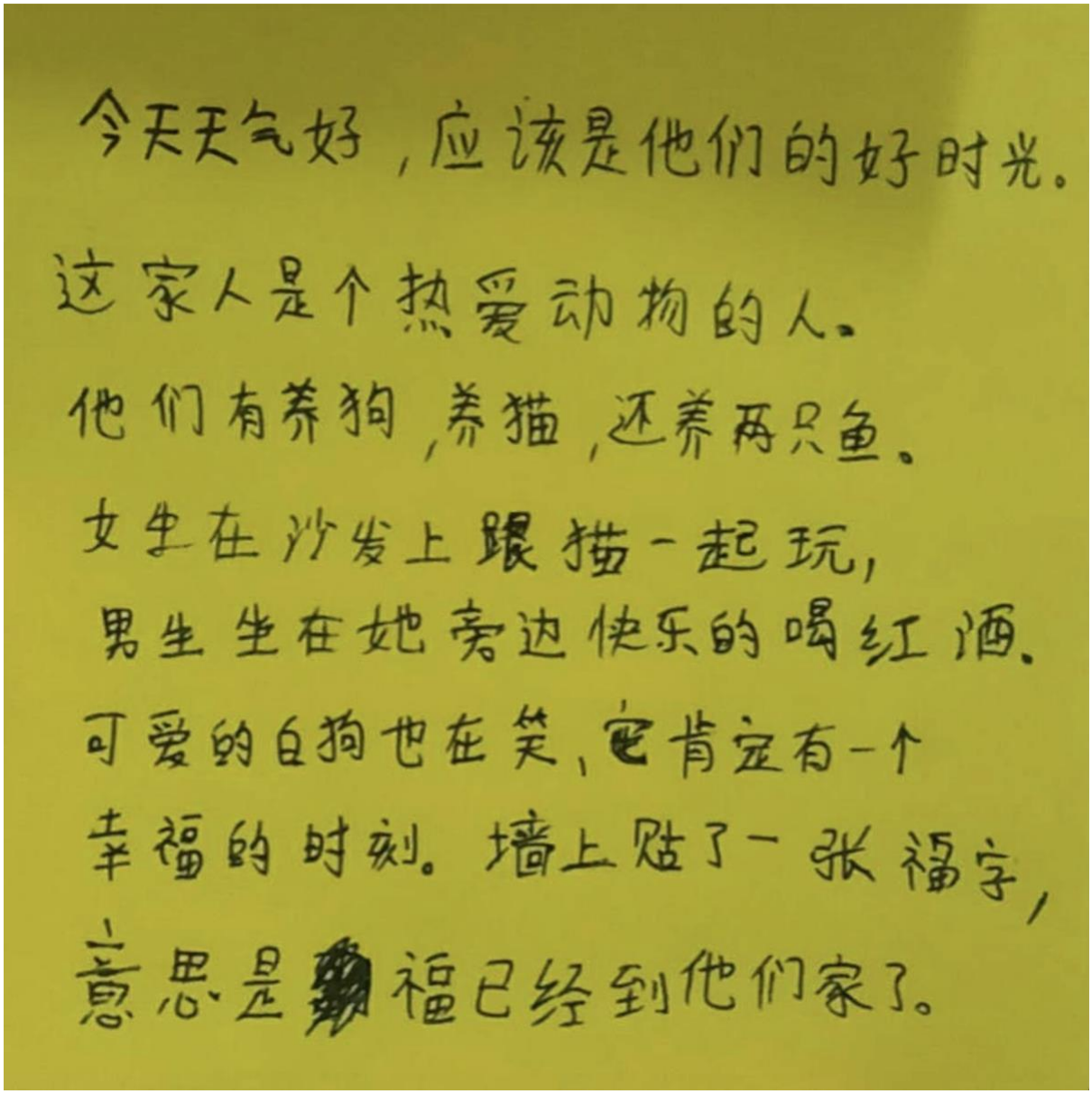

Figure 3: Yui's composition (Chinese).

Source: Author's photograph.

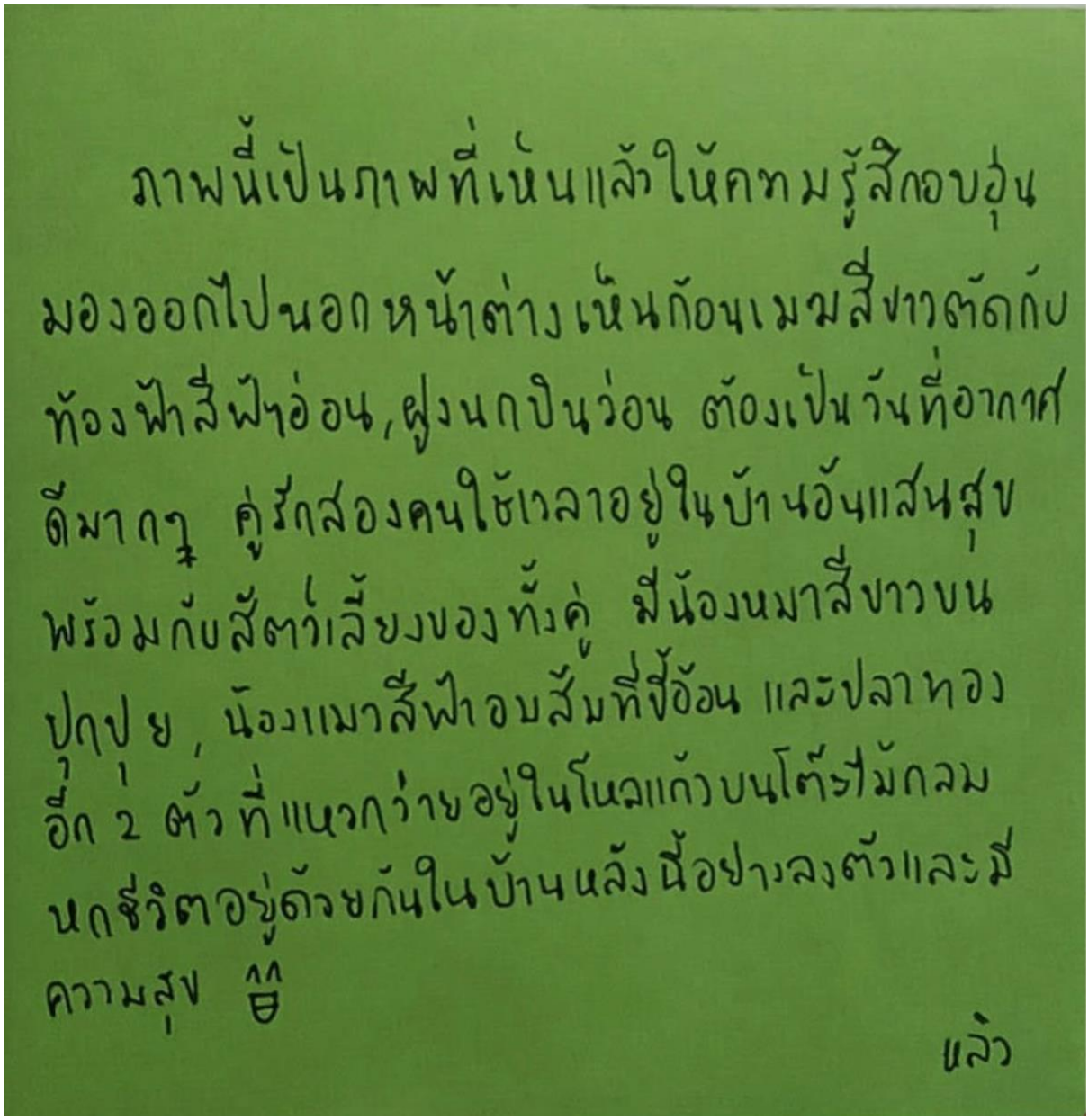

Figure 4: Yui's composition (Thai).

Source: Author's photograph.

Individual interview with Yui

1. Q: Why did you apply classifier inconsistently in the second sentence?

A: At first I wanna write 一只狗，一只猫，两条鱼 but I wrote the first stroke of the Character wrong, that is why …

2. Q: How do you learn classifiers?

A: I learnt some from my Chinese class. Watching Chinese series, listening to Chinese songs and talking to Chinese people do help me a lot in memorising the classifiers because when you hear something repeatedly, you will automatically remember it.

3. Q: Are classifiers difficult for you? Why?

A: I do not think it is hard, it is just complicated and takes time in memorising the correct one. In Thai, we also have lots of classifiers and each noun has its specific classifier as well, so I think it is the beauty of the language.

The hard cases for actualism: Efficacious vs inefficacious oughts

KIDA LIN

Abstract

Joe is a firefighter who is sent to rescue a child from a burning building. He is the only person on the scene, and he has a professional duty to do it. Unfortunately, he was too busy gambling to learn anything useful when he was at the firefighter school. As such, suppose that if he enters the burning building, it is sufficiently unlikely that he can successfully rescue the child. It is more probable he will die alongside the child. Ought Joe to save the child? In this paper, I argue that there are two distinct types of 'oughts'. In one sense, Joe ought to save the child and this is what I call the 'inefficacious' sense of oughts. In another sense, it is not the case that he ought to save the child, and this is the 'efficacious' sense of oughts. I show how this distinction helps actualism – the philosophical position that what matters for the assessment of options is what will actually happen – deal with a cluster of 'hard cases'.

Introduction

> **Weak-willed Susan**: Susan is meant to finish a group project today. It is part of her job, and she is capable of the work. But as it happens, she will not in fact do it; she will stay at home and watch TV. Nevertheless, there are two other things she can do. She can call her groupmates from home and let them know that she will not come to the office, in which case her groupmates will rush to do the work for her. Alternatively, she can simply do nothing. The best thing that can happen is if Susan does the project herself, followed by if she lets her groupmates know. The worst that can happen is if she chooses to do nothing. What ought Susan to do?

Philosophers known as 'actualists' argue that Susan ought to do the project herself, and, given that she will not, she ought to inform her groupmates. They contend that *what will in fact happen* is relevant to the assessment of options.[1] 'Possibilists' disagree, insisting that what matters for the assessment of options is *what is possible to happen*. It is possible for Susan to do the project herself

[1] Actualists: Sobel (1976), Goldman (1976), Jackson and Pargetter (1986), Jackson (2014).

and that is what she ought to do.[2] I suspect that most people's intuitions accord with actualism in this case. Of course, possibilists have presented multiple objections to the actualist answer. They argue that actualists commit us to an incompatible set of options (both for Susan to do the work herself and to ask her groupmates to do the work), which creates an 'action dilemma'.[3] They also argue that actualism lets the agent off the hook too easily, as actualism seems to allow for the fact that an agent will not do something to generate a less demanding option.[4] Relatedly, they charge that actualism does not take our agency seriously enough, and it is therefore agency-effacing.[5] These are all important objections, but I want to focus on a different concern in this paper: the concern that actualism leads to counterintuitive or even 'grotesque' conclusions in certain cases.[6] Consider the following example:

> **Absent-minded Michael**: Michael is looking after his granddaughter Emily, who asks him for a glass of water. As a matter of fact, Michael is about to accidentally give her a glass containing an arsenic solution. When he does so, she will drink the contents of the glass and die. Michael has many options besides giving Emily the arsenic. He could instead give her a glass of water as she requested. Or he could give her a glass of bleach. If he were to give her a glass of water, she would drink it and be happy. If, on the other hand, he was to give her a glass of bleach, then, while she would not drink enough of the liquid for it to be fatal, she would drink enough to suffer severe and irreversible damage to her body. What ought Michael to do?[7]

In this case, the actualist answer – that 'he ought to give her a glass of water, and given that he will not, he ought to give her a glass of bleach' – sounds much less plausible. Instead, most would side with possibilists and hold that Michael simply ought to give her a glass of water. We then have a further puzzle: how can these two otherwise identical cases yield different intuitions?

Let us call cases like Susan the *easy cases* for actualism and cases like Michael the *hard cases*. In the second section, I argue that the hard cases like Michael pose a problem for actualism. I do this by examining two responses that actualists may give to the hard cases. I then present and defend my own proposal. I argue that the key to differentiate hard cases from easy ones lie with differentiating the two types of 'oughts'.

---

[2] Possibilists: Goldman (1978), Feldman (1986), Zimmerman (2007). There are also those who try to strike a middle ground: Carlson (1999), Vorobej (2000), Bykvist (2002), Louise (2009), Ross (2012).

[3] Carlson (1999). For responses, see Jackson and Pargetter (1986), Jackson (2014).

[4] Carlson (1999). For responses, see Jackson and Pargetter (1986), Louise (2009).

[5] For discussion, see Louise (2009).

[6] Wedgwood (2009).

[7] Ross (2012), p. 75.

Before proceeding, three preliminary remarks are needed. First, I will use the label '*problem cases*' to refer to both the easy and the hard cases. Problem cases, as I stipulate, have the following structure: an agent S has three options φ1, φ2 and φ3. φ1 is the best option all-things-considered, followed by φ2, which is better than φ3. We assume that S will not φ1, and we then ask what she ought to do. When I use 'φ1, φ2, and φ3' below, I refer to 'the best option, the second-best option, and the worse option', respectively. Second, by 'S will not φ', I mean 'S φ-ing is sufficiently unlikely'. Similarly, 'S will φ' means 'S φ-ing is sufficiently likely'.[8] In addition, I will use 'oughts' and 'obligations' interchangeably in this paper. When I talk about 'what S ought to do', unless noted otherwise, I mean 'what S objectively ought to do'. I will further stipulate that what S objectively ought to do is just what she ought to do if she knows all the relevant facts about the case.[9] Finally, let us clarify what actualists and possibilists agree and disagree upon. They both agree that S ought to φ1 (namely, that Susan ought to do the project herself and Michael ought to give Emily a glass of water). What they disagree about is whether S ought to φ2. Only actualists hold that S ought to φ2. More controversially, I will assume that both sides agree that, 'S ought to φ2, *if* she will not φ1'.[10] What they disagree about is whether 'S ought to φ2, full stop'. To put it differently, I assume that actualists and possibilists agree on the conditional obligation that, if Susan will not do the project herself, she ought to inform her groupmates. However, only actualists hold that, *given* Susan will not do the project herself, she ought to inform her groupmates.[11]

Why do the hard cases pose a challenge for actualism?

Consider two responses that actualists may give to the hard cases. For one, actualists can bite the bullet and accept that Michael ought to give Emily a glass of bleach. To make this more palatable, they can emphasise the distinction between objective obligation and subjective prescription/deliberation. It will surely be bizarre if someone tells Michael that he ought to give his granddaughter a glass of bleach. It will also be inappropriate if Michael deliberates in a way that

[8] Louise (2009), p. 333.

[9] See Jackson (2014), pp. 640-43. Also Louise (2009).

[10] This is the point made in Jackson and Pargetter (1986).

[11] This assumption is controversial. For example, Kiesewetter (2015, pp. 934–36) would deny that possibilists accept 'S ought to φ2, *if* she won't φ1'. Instead, he holds that the possibilist answer is: S ought to {if she won't φ1, φ2}. The difference is that, according to Kiesewetter, the ought claims of possibilists are *wide scope* rather than *narrow scope*. Regardless, this complication wouldn't matter for my purposes in this paper.

involves considering giving his granddaughter a glass of bleach. Nevertheless, actualists can insist this is indeed what Michael objectively ought to do.

The problem with this response is that it won't convince many people who are puzzled by the hard cases. Many possibilists will still hold onto their conclusion that the actualist answer is counterintuitive, even after they recognise that it's the objective obligation we focus on. Consider an analogy. Some philosophers hold that we have very stringent obligations to help the global poor.[12] They often remind us that it is probably not pragmatically useful to tell people that they have such stringent obligations, as this will make them give up doing anything altogether. Nevertheless, for many of those who disagree, they often still hold onto their conviction, even when they agree that we are talking about the objective obligations and not about what's useful to tell other people.[13]

In addition, many possibilists take the hard cases like Michael as a decisive objection against actualism. Jacob Ross, for example, takes the case of Michael as showing that 'the main problem with actualism is that it's obviously false'.[14] Ralph Wedgwood holds that the hard cases render actualism 'a deeply objectionable view'.[15] Therefore, it seems all the more pressing that actualists can say something more to convince those who don't share their intuition.

Another way actualists can respond is to argue for a threshold on obligations. They can hold that 'S ought to φ' has a built-in requirement on how bad φ can be. One suggestion is that perhaps φ needs to be *pro tanto* permissible.[16] But this is an unacceptably ad hoc solution. The so-called 'lesser-evil' cases also show that sometimes the agent ought to do what is *pro tanto* impermissible to do.[17]

Another suggestion actualists can make is that perhaps φ needs to be permissible all-things-considered. That is, perhaps we ought only to do acts that are permissible all-things-considered. Even if this is true, it still begs the question. If actualists insist it is not true that Michael ought to give his granddaughter a glass of bleach because doing so is impermissible all-things-considered, we might want to know why this option is impermissible all-things-considered. More precisely, we

---

[12] E.g., Singer (1972), Sobel (2007).

[13] E.g., Kamm (2001, 2007), Hurka and Shubert (2012).

[14] Ross (2012), pp. 75–76.

[15] Wedgwood (2009).

[16] By *pro tanto* permissible acts, I just mean acts that are permissible absent countervailing reasons. For discussion on *pro tanto* obligations, and how they differ from prima facie obligations, see Reisner (2013).

[17] Lesser-evil cases are those where both options for the agent are impermissible. See, e.g., Lazar (2012, 2018).

want to know why doing the second-best option is impermissible all-things-considered in the hard cases, while doing the second-best option is not impermissible in the easy cases. That means we are back to the start – we need an account that distinguishes the easy cases from the hard ones.

Solving the hard cases: Efficacious and inefficacious oughts

We start with an observation that both the hard and the easy cases are cases where the agent will not do the best option (i.e., S won't φ1), and this seems to have a bearing on their obligations. After all, it is only because Susan will not do the project herself that leads actualists to contend that she then ought to tell her groupmates about it. But here is one difference between the hard and the easy cases: in the hard cases, we assume not only that the agent will not do the best option, but also that *she will do the worst option* (i.e., S will φ3). That is, in the case of Michael, we assume not just that he will not give Emily a glass of water, but also that he will give her a glass of an arsenic solution. Of course, S will φ3 entails that S won't φ1 – but, importantly, it also entails that S won't φ2. In other words, one feature that distinguishes the hard cases from the easy ones is that we stipulate in the hard cases that *the agent will not do the second-best option*. Given this structural difference, I posit the following hypothesis:

(1) In the problem cases, S ought to φ2 iff it's not the case that S won't φ2.[18]

Principle (1) vindicates our intuitive responses to the problem cases. In the easy cases like Susan, given it is not true that she will not inform her groupmates, informing her groupmates is therefore also something she ought to do (i.e., the actualist answer). In the hard cases like Michael, given he will not in fact give Emily a glass of bleach (as he will give her arsenic solution instead), giving her a glass of bleach is therefore *not* something he ought to do (i.e., the possibilist answer).

As I stipulated above, by 'S will not φ2', I mean 'S φ2-ing is sufficiently unlikely'. Note that 'it is not the case that S will not φ2' does not mean 'S will φ2'. That is, to say doing something is not sufficiently unlikely does not entail that doing it is sufficiently likely. Suppose you are indecisive between whether to have Indian or Chinese for dinner. It might be that it is not sufficiently unlikely that you choose Indian, nor is it sufficiently likely that you do so (you might say, 'it's 50-50'). This distinction is important because although it is not true that Susan will not inform her groupmates,

[18] This is a claim *about* the problem cases, and not a general claim that *figures into* the problem cases. That is, 'it's not the case that S won't φ2' is a sufficient and necessary condition for 'S ought to φ2' *only in the problem cases* – my claim is not that 'it's not the case that S won't φ2' is a sufficient and necessary condition *in general* which happens to apply to the problem cases. See, also note 20. I thank Prof. Southwood and an anonymous reviewer from the ANU Philosophy department for raising this point.

nor is it true that she will. Principle (1) holds that satisfying the former is enough for her to hold the obligation to inform her groupmates, and that this is why we can maintain the actualist answer in the easy cases.

Principle (1) gives us the intuitively right results, but can we motivate it with a coherent rationale, so as to avoid the worry that it is ad hoc? Consider first the more general worry about whether the likelihood of S φ-ing can have any bearing on S's obligations. It may be objected that our obligations have nothing to do with how likely it is for us to do certain things. But that cannot be true, because the actualist response in the easy cases depends on the idea that the likelihood for us to do certain things *does* have an impact on our obligations. After all, it is only because Susan will not do the project herself that prompts actualists to hold that she then ought to inform her groupmates.

A more specific worry possibilists might raise is why the fact that S will not φ2 blocks her obligation to φ2, but that S will not φ1 does not block her obligation to φ1. Namely, the central issue is why in the hard cases where the agent is sufficiently unlikely to either φ1 or φ2, she ought to φ1 but not φ2 (the possibilist answer, which we assume to be correct).

Here is my suggestion: the obligations to do φ1 and φ2 are distinct, and only the latter is sensitive to the likelihood of the agent φ-ing. To support this, I will argue first that there are two distinct types of 'oughts', and only one of them is sensitive to the likelihood of S φ-ing. I will then demonstrate why the obligation that S has to φ2 (in the easy cases) is this latter obligation.

To see whether there are two distinct types of obligations, consider the following:

> **Untrained firefighter**: Joe is a firefighter who is sent to rescue a child from a burning building. He is the only person on the scene, and he has a professional duty to do it. Unfortunately, he was too busy gambling to learn anything useful when he was at the firefighter school. As such, suppose that if he enters the burning building, it's sufficiently unlikely that he can successfully rescue the child. It is more probable that he will die alongside the child.

Ought Joe to save the child? I think we have conflicting intuitions here. In one sense, yes, he ought to save the child. That is because the child needs help, and Joe has the professional duty to rescue. But in another sense, no, it is not the case he ought to save the child – the reason being that it is sufficiently unlikely for him to succeed. If he enters the building, he will sacrifice his life without achieving any good.

Let us call the sense in which he ought to save the child the '*inefficacious*' sense of obligation, and the sense in which he ought not to rescue the child the '*efficacious*' sense.[19] I contend that we can distinguish inefficacious oughts from efficacious oughts as follows:

(2) S ought to φ, in an efficacious sense, only if it is not the case that S will not φ.[20]

Principle (2) states that efficacious oughts are conditional on it not being the case that the agent φ-ing is sufficiently unlikely, whereas inefficacious oughts have no such condition. This explains the conflicting intuitions in the case of Joe. Given that it is sufficiently unlikely for him to (successfully) rescue the child, it is not the case that he ought, in an efficacious sense, to rescue them. Nevertheless, he still ought, in an inefficacious sense, to rescue them. But what does that mean? Are there any other differences between efficacious and inefficacious oughts other than whether the ought claims are conditional on the likelihood of S φ-ing?

I think the following also holds:

(3) When S ought to φ, in an efficacious sense, S is compelled to take steps towards φ-ing. When S ought to φ, in an inefficacious sense, it is not necessarily the case that S is compelled to take steps towards φ-ing.

By 'take steps towards φ-ing', I mean 'doing what is necessary in order to φ'. For example, for Joe to 'take steps towards saving the child' is for him to enter the burning building, as entering the burning building is what is necessary to save the child. By 'S is compelled to X', I mean 'X is the appropriate response for S to have'. In this instance, I mean when S ought to φ in an efficacious sense, the appropriate response to this obligation is to take steps towards φ-ing.

Consider an analogy. Suppose you agreed to go to a music concert with me, and we agreed to buy our tickets separately in advance. Suppose that you did not. On the day when we meet outside the

---

[19] Following Southwood (2016, p. 17, n. 29), I shall remain neutral on whether our talk and thought of involving 'ought' are literally ambiguous or whether 'ought' claims can express different propositions without being literally ambiguous.

[20] Note that 'it's not the case that S won't φ' is only the *necessary* (but not sufficient) condition for S to hold an efficacious obligation. This is because there can be cases where it's not the case that S ought, in an efficacious sense, to do something, even if it's not sufficiently unlikely for him to succeed. For example, suppose instead that it's not sufficiently unlikely for Joe to rescue the child, but if he does try to rescue, it's sufficiently likely that he will suffer from severe burns. We might then argue that perhaps Joe ought not to rescue the child, in an efficacious sense, because the cost to him is too high (saving the child is now supererogatory). Some might also note that principle (2) only states the necessary condition for S's holding an efficacious obligation while principle (1) states the necessary and sufficient condition. This is not an issue for my account, as 'it's not the case that S won't φ2' is a necessary and sufficient condition for S to hold an efficacious obligation *in the problem cases*, although 'it's not the case that S won't φ' is only a necessary condition for S to hold an efficacious obligation *in general*. Some might also wonder whether the converse of principle (2) is true. Namely, some might wonder whether 'S ought to φ, in an inefficacious sense, if it's the case that S won't φ'. I don't think so. There are cases where the agent can't φ and she is not culpable for the fact that she can't φ. In those cases, it's not true that she ought to φ, in an inefficacious sense. See my discussion below.

concert hall, I might say to you, 'you ought to go to the concert with me!' What I am referring to in this instance is an inefficacious obligation, because the appropriate response for you to have is not to do what is necessary to go to the concert with me. Entering the concert hall, for example, seems to be what is necessary for you to go to the concert with me. Clearly, the appropriate response for you to have when I complain to you is clearly not to enter the concert hall without a ticket.

Principle (3) explains what is entailed when we say, 'in one sense, Joe ought to save the child; in another sense, it is not the case that he ought to do so'. What we mean is that Joe ought to save the child, although he should not enter the burning building. This indicates that the obligation Joe has is an inefficacious one.[21]

Before we move on to consider how efficacious and inefficacious oughts apply to the problem cases discussed above, consider two objections to the case of Joe. Some might reject that Joe ought to save the child. They might argue this along the line of 'ought implies can' principle.[22] That is, Joe ought to save the child only if he can save the child, in a relevant sense. Because he cannot save the child, it is not the case that he ought to. Despite all the disagreement around whether ought does imply can, I think most would agree that if S is culpable for the fact that she cannot φ, this does not block her obligation to φ. I ought to submit this paper tonight. I cannot do that if I destroy my laptop now. Surely, I still ought to do so even if I do destroy my laptop, and thereby making myself incapable of submitting the paper. Similarly, the fact that Joe cannot save the child does not block his obligation to do so, because he is culpable for the incapability.

Some might then deny that there is any sense in which it is not true that Joe ought to save the child. They might argue that because Joe is the only one culpable here, he unequivocally ought to save the child. But again, when we have an obligation to do something, it often means that we have an obligation to do what is necessary to achieve that thing. I ought to submit this paper tonight, and writing it now is what is necessary for me to submit the paper. I therefore ought to write it now. By contrast, it is implausible to maintain that Joe ought to go into the burning building, despite it being necessary for fulfilling his obligation to save the child. Therefore, it seems that there is indeed a sense in which it is not the case that Joe ought to save the child.

---

[21] Kiesewetter (2015) defends what he calls the '*transmission principle*', which holds that 'if you ought to perform a certain act, and some other action is a necessary means for you to perform that act, then you ought to perform that other action as well'. Principle (3) can therefore be seen as holding that *transmission principle only applies to efficacious oughts but not inefficacious oughts*.

[22] For discussion of 'ought implies can', see Mason (2004), Vranas (2007), Cohen (2008), Gheaus (2013) and Estlund (2014).

Let us now turn back to the problem cases.[23] In the easy cases, S ought to φ1, and given that she will not, she ought to φ2. In the hard cases, S ought to φ1, but it is not the case that she ought to φ2. We speculated above that this difference in obligations exists because in the hard cases, S will not φ2 (i.e., S φ2-ing is sufficiently unlikely). Now we have an explanation as to why the fact that S will not φ2 in the problem cases means that it is not the case that she ought to φ2 (principle (1)): the sense in which she ought to φ2 is the efficacious sense. Because S ought to do something in an efficacious sense only if it is not true that she will not do that thing (principle (2)), and because it is true that she will not do φ2 in the hard cases, S ought not to φ2 in the hard cases.

But a question remains: why should we believe that the sense in which the agent ought to φ2 in the problem cases (as opposed to the sense in which she ought to φ1) is the efficacious sense? The answer: because the appropriate response for S to have in virtue of her obligation to φ2 is to do what is necessary to φ2. The appropriate response for S to have in virtue of her obligation to φ1 is not, however, to do what is necessary to φ1 (principle (3)). Consider:

> **Professor Procrastinate**: Procrastinate receives an invitation to review a book. He is the best person to do the review and has the time. The best thing that can happen is that he says yes, and then writes the review. However, were he to say yes, he would not in fact get around to writing the review – this is the worst that can happen. The second-best thing is if he says no and Dr Reliable writes the review instead.[24]

In this classic (easy) case, most agree that Professor Procrastinate ought to accept the invitation and do the review (φ1), and given that he will not, he ought to say no (φ2). Why is the sense in which he ought to accept the invitation and do the review the inefficacious sense? Because this obligation does not entail that Procrastinate ought to do what is necessary to fulfil it – namely, to accept the invitation. Otherwise, it would indeed create an 'action dilemma' (i.e., he ought to accept and reject the invitation). Similarly, the sense in which he ought to say no is the efficacious sense, as it does follow from this that Professor Procrastinate ought to do what is necessary to fulfil the obligation (perhaps, by emailing the editor to turn down the invitation).

To sum up my proposal: in the problem cases, the sense in which the agent ought to do the best option is the inefficacious sense, and the sense in which she ought to do the second-best is the efficacious sense. For the agent to hold efficacious obligations, it must not be the case that her fulfilling the obligation is sufficiently unlikely. The hard cases differ from the easy ones in that the

---

[23] To avoid confusion: the case of Joe is not an easy or a hard case. I introduce it in order to demonstrate and make plausible the distinction between efficacious and inefficacious oughts.

[24] Jackson and Pargetter (1986).

agent is sufficiently unlikely to do the second-best. Therefore, it is not true that the agent ought to do the second-best in the hard cases, despite the fact that she ought to do so in the easy cases.[25]

Now, consider an implication of my proposal:

> **Irresponsible Lucy**: Lucy owes Katie $500. She has made the promise to pay back the money today. As it happens, she does not actually have $500 now, as she has spent all her money on the World Cup. Alternatively, she can pay Katie $300, which is not as good as repaying in full, but is certainly better than nothing. Unfortunately, Lucy does not have $300 either. What ought Lucy to do?

I suspect most would agree that what she ought to do is to pay back $500. It would not make much sense to say, 'she ought to pay $500, and given that she will not, she ought to pay $300'. Such talk only makes sense in the context where her repaying $300 is not sufficiently unlikely. Therefore, it seems that we can apply the distinction between efficacious and inefficacious oughts to account for this case – she ought, in an inefficacious sense, to pay $500. It is *not* the case that she ought, in an efficacious sense, to pay $300 (given that she will not pay $300), because doing so is sufficiently unlikely.

But this case differs from the hard cases like Michael in one respect. In cases like Michael's, the agent will not do what is best (i.e., S will not φ1) but she nonetheless *can* do what is best, in the relevant sense. In the case of Lucy, however, she not only will not do what is best (i.e., repaying $500), she also *cannot* do it (i.e., she does not have the money now). Regardless, as our discussion of firefighter Joe suggests, the fact that the agent cannot do something when she is culpable for that incapacity does not block her inefficacious obligation to do that thing. As Lucy cannot repay the

---

[25] In a sense, my argument can be understood as supporting a particular formulation of actualism. That is, there are two ways to formulate actualism: a) for any options φ1, φ2, φ3, such that φ1 is morally preferable to φ2 and φ2 is morally preferable to φ3, if S will not φ1, S ought to φ2; b) out of any set of options {φ1 … φn}, S ought to do the best remaining option after excluding all options that she will not in fact do. My argument is therefore an argument about why b) is the more plausible version of actualism. I thank an anonymous reviewer from the ANU Philosophy department for this point. I agree that my argument does support b) as a more plausible formulation, but I will note that it is not always clear which formulation actualists adopt, and this may explain why the debate about the hard cases can be confusing. For example, Jackson and Pargetter (1986, p. 233) state that '[b]y Actualism we will mean the view that the values that should figure in determining which option is the best and so ought to be done out of a set of options are the values of what would be the case were the agent to adopt or carry out the option, where what would be the case includes of course what the agent would simultaneously or subsequently in fact do: the (relevant) value of an option is the value of what would in fact be the case were the agent to perform it'. This is not an explicit endorsement of either a) or b), as it only tells us *how we should determine the value of an option*, φ – we determine it by looking at what will actually happen – and not *when an option ought to be chosen*. As such, before presenting a hard case cited above as an objection, Ross (2012, p. 76) formulates actualism as the view that 'an agent ought to ϕ just in case she ought to prefer the outcome that would result from her ϕ-ing to the outcome of that would result from her not ϕ-ing'. This is a view about not only how we should determine the value of an option but also when an option ought to be chosen (i.e., it ought to be chosen just in case its alternative, φ', is better—which is close to though not identical with formulation a)). As such, my argument serves to clarify the dispute.

money only because she has spent all her money, she ought, in an inefficacious sense, to repay $500.

The implication of this is that we can now broaden the types of cases relevant to the actualism–possibilism debate. Traditionally, cases that are relevant to the actualism–possibilism debate are only those where S will not φ1 but can still φ1. Our discussion suggests that cases like Lucy's, where S will not and cannot φ1, and she is culpable for the incapacity, are also relevant to the actualism–possibilism debate.[26]

Conclusion

In this paper, I have discussed what I call the 'hard cases' for actualism. These are cases where S ought to φ1, but it is not the case that S ought to φ2. They differ from the easy cases for actualism, where the actualist answer is the intuitive one. I have first argued that the hard cases indeed pose a problem for actualism, as the actualists' attempts to bite the bullet or impose a threshold are not satisfactory ways to address these cases. I then put forward my account, which holds that the sense in which S ought to φ2 in the problem cases is the efficacious sense. Because S ought to φ, in an efficacious sense, only if it is not true that S will not φ, and because it is true that S will not φ2 in the hard cases, it is not the case that S ought to φ2 in the hard cases. In other words, Susan ought, in an efficacious sense, to inform her groupmates because it is not sufficiently unlikely that she does so. By contrast, it is not the case that Michael ought, in an efficacious sense, to give his granddaughter a glass of bleach as it is indeed sufficiently unlikely for him to do so.

Acknowledgements

I would like to thank Associate Professors Nicholas Southwood and Seth Lazar for their comments.

[26] Some may object that my proposal is not an actualist proposal after all. This is because it may be argued that actualists must insist one only ought to do the second best (i.e., φ2) if one is robustly disposed to fail to do the best option (i.e., φ1), and it seems that my examples like Susan are not really cases where the agent is robustly disposed to fail to do the best option. As such, it doesn't seem to vindicate actualism. In reply: I don't think actualism only applies to cases where the agent is robustly disposed to fail to φ1 (at least, not if 'robustly disposed to fail to φ1' means anything more than 'will not/sufficiently unlikely to φ1'). Suppose Professor Procrastinate would have been able to stop procrastinating if he took a magic pill yesterday. Suppose that he did not, in which case I still don't think he should say yes to the invitation (i.e., an actualist answer). This is so even if it was really quite likely that he took the pill yesterday, so his tendency to procrastinate is not modally robust. I think modal robustness only matters to the extent that it means Professor ought to φ1 in an inefficacious sense. That is, if the tendency to procrastinate is modally robust in that there was nothing that Procrastinate could have done to stop procrastinating, perhaps it is not the case that Procrastinate ought to accept the invitation and do the review (or at least he is excused from failing to accept and review – this may be so because the obligation to accept and review is generated from sources other than his ability, such as his professional duty or his promise). I thank Prof. Southwood for this objection.

References

Bykvist, K, 2002. Alternative actions and the spirit of consequentialism. *Philosophical Studies*, *107*(1), pp. 45–68. doi.org/10.1023/A:1013191909430.

Carlson, E, 1999. Consequentialism, alternatives, and actualism. *Philosophical Studies*, *96*(3), pp. 253–68. doi.org/10.1023/A:1004239306956.

Cohen, GA, 2008. *Rescuing justice and equality*. Harvard University Press. doi.org/10.4159/9780674029651.

Estlund, D, 2014. Utopophobia. *Philosophy & Public Affairs*, *42*(2), pp. 113–34. doi.org/10.1111/papa.12031.

Feldman, F, 1986. *Doing the best we can: An essay in informal deontic logic* (Vol. 35). Springer Science & Business Media. doi.org/10.1007/978-94-009-4570-8.

Gheaus, A, 2013. The feasibility constraint on the concept of justice. *The Philosophical Quarterly*, *63*(252), pp. 445–64. doi.org/10.1111/1467-9213.12058.

Goldman, HS, 1976. Dated rightness and moral imperfection. *The Philosophical Review*, *85*(4), pp. 449–87. doi.org/10.2307/2184275.

Goldman, HS, 1978. Doing the best one can. In *Values and morals* (pp. 185–214). Springer, Dordrecht. doi.org/10.1007/978-94-015-7634-5_11.

Hurka, T, and Shubert, E, 2012. Permissions to do less than the best: A moving band. *Oxford studies in normative ethics*, *2*, pp. 1–27.

Jackson, F, 2014. Procrastinate revisited. *Pacific Philosophical Quarterly*, *95*(4), pp. 634–47. doi.org/10.1111/papq.12051.

Jackson, F, & Pargetter, R, 1986. Oughts, options, and actualism. *Philosophical Review* 95, pp. 233–55. doi.org/10.2307/2185591.

Kamm, FM, 2001. *Morality, mortality: Rights, duties, and status* (Vol. 2). Oxford University Press on Demand. doi.org/10.1093/0195144023.001.0001.

Kamm, FM, 2007. *Intricate ethics: Rights, responsibilities, and permissible harm*. OUP USA. doi.org/10.1093/acprof:oso/9780195189698.001.0001.

Kiesewetter, B, 2015. Instrumental normativity: In defense of the transmission principle. *Ethics*, *125*(4), pp. 921–46. doi.org/10.1086/680911.

Lazar, S, 2012. Necessity in self-defense and war. *Philosophy & Public Affairs*, *40*(1), pp. 3–44. doi.org/10.1111/j.1088-4963.2012.01214.x.

Lazar, S, 2018. In dubious battle: uncertainty and the ethics of killing. *Philosophical Studies*, *175*(4), pp. 859–83. doi.org/10.1007/s11098-017-0896-3.

Louise, J, 2009. I won't do it! Self-prediction, moral obligation and moral deliberation. *Philosophical Studies*, *146*(3), pp. 327–48. doi.org/10.1007/s11098-008-9258-5.

Mason, A, 2004. Just constraints. *British Journal of Political Science*, *34*(2), pp. 251–68. doi.org/10.1017/S0007123404000043.

Reisner, AE, 2013. Prima facie and pro tanto oughts. *International Encyclopedia of Ethics*. doi.org/10.1002/9781444367072.wbiee406.

Ross, J, 2012. Actualism, possibilism, and beyond. *Oxford Studies in Normative Ethics*, *2*, pp. 74–96. doi.org/10.1093/acprof:oso/9780199662951.003.0004.

Singer, P, 1972. Famine, affluence, and morality. *Philosophy & Public Affairs*, pp. 229–43.

Sobel, JH, 1976. Utilitarianism and past and future mistakes. *Noûs*, pp. 195–219. doi.org/10.2307/2214575.

Sobel, D, 2007. The impotence of the demandingness objection. *Philosophers' Imprint*, *7*(8), pp. 1–17.

Southwood, N, 2016. Does 'Ought' Imply 'Feasible'? *Philosophy & Public Affairs*, *44*(1), pp. 7–45. doi.org/10.1111/papa.12067.

Vorobej, M, 2000. Prosaic possibilism. *Philosophical Studies*, *97*(2), pp. 131–36. doi.org/10.1023/A:1008349904432.

Vranas, PB, 2007. I ought, therefore I can. *Philosophical studies*, *136*(2), pp. 167–216. doi.org/10.1007/s11098-007-9071-6.

Wedgwood, R, 2009. 'Against actualism', *PEA Soup*, posted on September 11, 2009. peasoup.typepad.com/peasoup/2009/09/against-actualism.html.

Zimmerman, MJ, 2007. *The concept of moral obligation*. Cambridge University Press.

'We make angels': Rediscovering the Victorian 'Angel in the House' in Spike Jonze's *Her* (2013) and Denis Villeneuve's *Blade Runner 2049* (2017)

ANNIKA MORLING

Abstract

Spike Jonze's 2013 film *Her* and Denis Villeneuve's 2017 film *Blade Runner 2049* (*BR2049*) both present spectral female characters who, despite being wholly constructed by futuristic technology, echo the Victorian image of the 'Angel in the House'. Just as male-engineered changes in technology altered women's roles and responsibilities in the Victorian period, both *Her* and *BR2049* show the software of their respective female characters to be designed by men. Just as the Victorian 'Angel' was perceived as belonging to the domestic sphere, yet denied ownership of property, the technological women of *Her* and *BR2049* are shown to be literally contained within masculine structures, with little personal autonomy. The technologically constructed women are also shown to possess a diminished corporeality, which echoes Victorian conceptions of ideal womanhood as spectral and omnipresent. Lastly, like the ideal Victorian 'Angel', both technological women are shown to exert an ameliorative moral influence over their respective male partners. While they present speculative visions of futuristic societies, both *Her* and *BR2049* paradoxically engage with Victorian ideas of the Angel in the House, suggesting an iterated pattern of female experience related to rapid technological change.

Introduction

The Victorian ideal of the 'Angel in the House' – an ornamental, salutary housewife of the upper middle class – emerged from a period of technological advancement in Britain over the late eighteenth and early nineteenth centuries. A product of literature, political rhetoric and societal gender norms, the 'Angel' ostensibly possessed a diminished corporeality, allowing her to imbue the domestic space with a calming feminine influence. She also denied her self-development in order to provide stability and reformation for her family in a time of economic and social change. In this essay I suggest that

these characteristics are reflected, somewhat paradoxically, in the female-coded artificial intelligence of two recent speculative fiction films – Spike Jonze's 2013 film *Her*[1] and Denis Villeneuve's 2017 film *Blade Runner 2049* (*BR2049*).[2] *Her*'s Samantha, an artificially intelligent computer operating system, and *BR2049*'s Joi, a hologram, are each the product of male-dominated technological processes, and are constructed to perform an ideal form of womanhood which reflects that of Victorian times. Echoing the confinement of the Victorian Angel to the domestic sphere, Samantha and Joi are both pieces of software contained within masculine-coded hardware. Both women are also bodiless, and this spectrality allows them to constantly recreate the comforts of the domestic space for their respective male partners. Lastly, both women are configured to devote themselves wholly to the development and stabilisation of their male partners. The re-emergence of the Angel in the House motif in the futuristic world of the speculative fiction film suggests how technological change, pioneered predominantly by men, alters societal conceptions of normative womanhood, and speaks to iterated patterns of female experience associated with periods of rapid technological advancement over time.

Gender and technology

The Victorian era was a period of significant technological and economic development in Britain. During this time, scientific advancements, the implementation of novel and transformative social policies, and the increased mechanisation of modes of production all transformed British society and its dominant ideologies across class boundaries. On a social level, this transformation altered previously held conceptions of national and personal identity, particularly in regards to class and gender. Increasing industrialisation over the late eighteenth century had rapidly expanded the British middle classes; large numbers of men and women moved from rural areas to urban centres, reflecting the broader transformation of Britain from a largely agrarian to an industrial economy.[3] Yet women were largely excluded from the productive aspects of this economy. It was widely understood that a middle-class wife did not partake in wage labour and was instead supposed to undertake philanthropic and household work.[4] Industrialisation thus created a large number of middle-class women who were not of noble birth, yet neither worked in the fields nor in the factories. Given their removal from

[1] *Her*, directed by Spike Jonze (Culver City: Sony Pictures Home Entertainment, 2013). DVD.

[2] *Blade Runner 2049*, directed by Denis Villeneuve (Culver City: Sony Pictures Home Entertainment, 2017). DVD.

[3] Joan Perkin, *Victorian Women* (London: John Murray, 1993), 87.

[4] It should be noted that women of the working classes constituted a significant proportion of workers in the farming, textile and proto-industrial industries in the early nineteenth century, and thus made a considerable contribution to production in the Victorian era. This essay focuses on perceptions of the upper-middle class Victorian woman. See: Antony Wrigley, *Continuity, Chance and Change: The Character of the Industrial Revolution in England* (Cambridge: Cambridge University Press, 1990), 85–7; Maxine Berg, 'What Difference did Women's Work Make to the Industrial Revolution?' *History Workshop* 35 (1993): 22–44.

productive processes, the Victorian era saw the emergence of the 'woman question', in which expected societal roles for women were recurrently discussed and redefined.[5] A popular conception of the ideal middle-class wife, promulgated predominantly by men, came to be the 'Angel in the House' – a 'decoratively idle' figure[6] who exuded a stabilising moral influence within her household.[7]

The technological and scientific advancements which so altered conceptions of womanhood in the Victorian era, and which thereby gave rise to the image of the Angel in the House, were engineered overwhelmingly by men. Prominent figures of British industrialisation included, for example, inventor James Watt, philosopher Jeremy Bentham and industrialist Abraham Darby.[8] Victorian women of all classes were systemically barred from such 'power-wielding occupations' as medicine, law, philosophy and science.[9] The systemic male dominance of production and technology in the Victorian era accords with contemporary feminist theories regarding the relationship between science, gender, and social power. If the aim of science, as Gill Kirkup suggests, is to attain 'power over the material world through knowledge', it follows that the removal of women from scientific knowledge results in their marginalisation in the material world.[10] In a similar vein, gender theorist Cynthia Cockburn contends that men maintain privilege and power in a society primarily through their control of technology.[11] Jonze and Villeneuve engage with these ideas in their respective speculative fiction films; in presenting female characters who are literally constructed by technology, *Her* and *BR2049* explore how female subjectivity and normative social roles are destabilised and subsequently reconstructed by male-controlled technological change.

Villeneuve's *BR2049* explicitly portrays technological and productive processes as extensions of male social power. In Ridley Scott's original film *Blade Runner* (1982), futuristic Los Angeles is under the control of the imperious Eldon Tyrell, a man who pioneered the large-scale production of highly realistic androids known as replicants and thereby drastically altered the physical and social landscape of the city. The Los Angeles of Villeneuve's 2017 sequel is controlled by Niander Wallace, a similarly autocratic figure who, after making his fortune in synthetic farming, bought the declining Tyrell Corporation and assumed control over its production of replicants. This transferral of power

---

[5] Diana Cordea, 'Two Approaches on the Philosophy of Separate Spheres in Mid-Victorian England: John Ruskin and John Stuart Mill'. *Procedia: Social and Behavioural Sciences* 71 (2013): 116.

[6] Perkin, *Victorian Women*, 86.

[7] The extent to which women actually resembled the ideal 'Angel in the House' image is subject to debate – see Mildred Jeanne Peterson, *Family, Love, and Work in the Lives of Victorian Gentlewomen* (Bloomington: Indiana University Press, 1989). Peterson claims that 'we have mistaken Victorian rhetoric for reality' in historical studies of upper-class Victorian women.

[8] Vanessa Dickerson, *Victorian Ghosts in the Noontide* (Columbia: University of Missouri Press, 1996), 4.

[9] Dickerson, *Victorian Ghosts*, 5.

[10] Gill Kirkup, 'Introduction', in *The Gendered Cyborg: A Reader*, ed. Gill Kirkup et al (London: Routledge, 2000), 3.

[11] Cynthia Cockburn, *Machinery of Dominance: Women, Men and Technical Know-How* (London: Pluto Press, 1988), 6–9, 33–4.

from one corporate, patriarchal regime to another is reminiscent of the British aristocratic practice of primogeniture, in which property and capital were consolidated within families through strictly male-line inheritance, precluding women's ownership over place. Like a great house of the Victorian era, then, the city of *BR2049* can be seen as a masculine dynastic space, in which women are removed from power.

In accordance with Cockburn's theory, Wallace's command over technology also facilitates his influence over dynamics of gender. His laconic proclamation to his assistant Luv, 'we make angels', can be interpreted as referring not only to his literal production of replicants but also to his ability to construct and alter gendered subject positions, especially regarding women. When Wallace examines his latest model of replicant, the naked woman slides out of plastic wrapping hanging from the ceiling and falls at Wallace's feet, symbolising Wallace's power to mould a new image of ideal womanhood. It is significant, therefore, that the Wallace Corporation is also responsible for the development of the Joi program, an iteration of which constitutes the film's primary female character. A colourful female-presenting hologram, Joi offers an attractive illusion of companionship and gratification against a backdrop of grey metropolitan dystopia – she is, to use Wallace's own term, engineered to be an 'angel' of the domestic space. Every time Joi is deactivated by her male owner and 'partner' K, the film's protagonist, her body-space is replaced by the Wallace Industries insignia. This effectively communicates how her containment within masculine power structures and her diminished corporeality – in short, her alignment with Victorian ideals of womanhood – are due to male-controlled technology.

While *BR2049* presents Niander Wallace as a key figure behind technological change, the narrative of Spike Jonze's *Her* is tightly focalised on its protagonist Theodore, and thus does not offer explicit insight into broader processes of technological production. That being said, the advertisement for the OS1[12] witnessed by Theodore in the railway station is voiced by a man, implying the off-screen male domination of software production. Moreover, this brief advertisement scene emphasises how Samantha is ultimately a product whose characteristics, including her ability to partake in a heterosexual relationship, are configured by presumably male software engineers. As Andrea Virginas argues, 'Theodore's spontaneous reaction in *Her* configuring his operating system as female suggests that AIs are brought into existence by humans' (male scientists') intellectual curiosity and emotional needs regulated by a heteronormative framework.'[13] Samantha and Joi are both literal products in masculinised, neoliberal marketplaces, and are hence engineered to perform an idealised form of womanhood. Advertised as an 'intuitive entity that listens to you, understands you, and knows you',

[12] The OS1 is the world's first artificially intelligent operating system, of which Samantha is a version.

[13] Andrea Virginas, 'Gendered Transmediation of the Digital from *S1m0ne* to *Ex Machina*: "Visual Pleasure" Reloaded?' *European Journal of English Studies* 21, no. 3 (2017): 298.

Samantha is, as Troy Bordun notes, 'the ideal woman for unrealistic male fantasies'.[14] Similarly, fluorescent advertisements for Joi in *BR2049* promise the anonymous (male, heterosexual) consumer 'everything [they] want'. The digital women of *Her* and *BR2049* can thus be seen as occupying a comparable position to the Victorian Angel in the House in that they too represent an idealised conception of womanhood constructed by men, within the context of rapid technological development. In a further parallel with Victorian gender norms, a significant characteristic of this engineered ideal womanhood is that it is dependent upon masculine structures and confined to the sheltered domestic space.

Containment within masculine structures

The digital women in *Her* and *BR2049* further reflect the Victorian Angel in the House image in that they are contained, due to their programming, within masculine structures of power. As Diana Cordea has explained in her comparative study of John Ruskin and John Stuart Mill, the Angel in the House figure reflected and was supported by contemporary Victorian conceptions of separate social spheres based on supposed biological gender difference.[15] Separate spheres ideology, as espoused by prominent social thinker John Ruskin in an 1864 speech, held that man, as the 'doer, the creator, the discoverer and the defender', naturally belonged within the public and productive sphere of society.[16] Women, on the other hand, possessed an intellect naturally disposed towards 'sweet ordering, arrangement and decision', and thus belonged within the private, domestic sphere.[17] Such notions of public and private separation confined women to an 'ancillary circle of housewifely and philanthropic activity', while men moved laterally, the transformers and producers of the material world.[18]

In addition to her rhetorical containment to the domestic sphere, the Victorian woman was in a more literal sense contained within masculinised physical structures due to property laws. Prior to the passage of the second Married Woman's Property Act in 1882, the legal doctrine of coverture held that a British woman's personal property, and any property she might subsequently acquire, legally passed to her husband upon her marriage.[19] It was thought that allowing married women equivalent property rights to men would cause, in the words of House of Commons member Philip Muntz, 'great

---

[14] Troy Bordun, 'On the Off-Screen Voice: Sound and Vision in Spike Jonze's *Her'*. *Cineaction* 98 (2016): 59.

[15] Cordea, 'Two Approaches', 118.

[16] John Ruskin quoted in Kate Millet, 'The Debate Over Women: Ruskin vs Mill', in *Suffer and Be Still: Women in the Victorian Age,* ed. Martha Vicinus (Bloomington: Indiana University Press, 1972), 126.

[17] Ibid.

[18] Millet, 'The Debate Over Women', 134.

[19] Ben Griffin, 'Class, Gender, and Liberalism in Parliament, 1868–1882: The Case of the Married Women's Property Acts', *Historical Review* 46, no. 1 (2003): 62.

difficulties in all the domestic arrangements of life'.[20] Such an objection reveals that separate spheres ideology rested on not only the idea of natural female command over the domestic sphere, but also on the idea that women's occupation of this sphere must be circumscribed by resolutely masculine power.

This idea is particularly explored in *BR2049*; Joi reflects idealised Victorian femininity in that she is literally contained within a masculine space, and relies upon that space for survival. In a clear allusion to contained female domesticity, Joi first materialises out of a projector fixed to the ceiling of K's apartment; her range of movement is constrained by the limited reach of the projector's arm, and as such, her holographic form cannot move beyond his living room. When Joi tells K that she is 'getting cabin fever' due to these conditions, she echoes the voice of the Victorian woman restricted by male rhetoric to a separate female sphere; the woman who was, 'by her office and place … protected from all [the] danger and temptation' of the outside world.[21]

Joi's complaint is supposedly alleviated by K's purchase of an 'emanator', a small, supplementary technological device sold by Wallace Industries which allows Joi's body to coalesce without a fixed light source; she may now go, as K tells her, 'anywhere [she] want[s] in the world now'. Yet the emanator, held within K's pocket, is itself a masculinised space, one which contains Joi until K chooses to summon her into existence. When the emanator is vindictively crushed by Luv in the closing stages of the film, Joi permanently vanishes; her concurrent 'death' with the destruction of the emanator reinforces her lack of independent subjectivity in the material world, and her reliance on male structures of power for existence. Joi's containment within the male-coded walls of K's apartment parallels the image of protection and containment inherent in the Angel in the House ideal. This suggests an iterated pattern of ideal womanhood as contained, protected and removed from productive societal spheres.

A diminished corporeality

In her ostensibly natural command over the domestic sphere, the ideal Victorian woman was perceived as possessing a diminished corporeality. Removed from productive modes and insuperably distant from the material conditions of life, the Victorian woman became, in public consciousness, 'equivocal, ambiguous, marginal, ghostly'.[22] This sympathy between women and ghosts is significant given the emergent interest in spiritualism in the Victorian era, a trend which developed alongside and in opposition to the growing 'masculine' fields of science and technology. In their ostensible

[20] Philip Muntz quoted in Griffin, 'Class, Gender and Liberalism in Parliament', 62.

[21] Ruskin quoted in Millet, 'The Debate Over Women', 126.

[22] Dickerson, *Victorian Ghosts*, 4–5.

possession of 'finer instincts'[23] and their natural authority in the fields of emotion and morality, women were perceived as possessing an inherent sympathy with the immaterial, supernatural realm. Literature of the time, written by both men and women, reflected the alignment of womanhood and the ethereal in British popular consciousness. Coventry Patmore's 1864 narrative poem 'The Angel in the House', after which the ideal Victorian woman is named, offers one example of this alignment: Patmore describes his wife as possessing a 'countenance angelical', which produces 'phantasms as absurd and sweet'; he claims her beauty 'haunts him all the night'; and reflects that 'everywhere I seem'd to meet/The haunting fairness of her face'.[24] Patmore's 'Angel' possesses substantial influence within the domestic space, yet this influence is crucially divorced from her physical form.

While male writers presented women as spectral and ethereal as part of an idealised vision of feminine domesticity, women did so as a means to 'create public discourse for voicing feminine concerns'.[25] As Vanessa Dickerson notes in *Victorian Ghosts in the Noontide*, female writers such as Elizabeth Gaskell, Vernon Lee and Christina Rossetti frequently adopted motifs of the supernatural and the ghostly in their work so as to self-reflexively explore affinities between ghosts and the female condition, such as their shared removal from productive and social power.[26] Christina Rossetti's 1864 poem 'After Death' can be read as one such work; featuring a female narrative voice speaking from beyond the grave, the poem embodies the liminality of the Victorian woman positioned between the physical and immaterial, or spiritual, worlds:

> He leaned above me, thinking that I slept
>
> And could not hear him; but I heard him say,
>
> 'Poor child, poor child': and as he turned away
>
> Came a deep silence, and I knew he wept …[27]

Of particular note in this poem is the ambiguous and pervasive power of its subject; while the woman is denied physical power due to her lack of 'body', Rossetti ascribes a pervasive perceptive ability to her bodiless form; the female spirit 'hears' the man, and 'knows' that he weeps. In this way, Rossetti's poem reflects the quasi-supernatural powers ostensibly wielded by the married Victorian woman. In her diminished corporeality, the Angel in the House was seen as occupying a perceptive, near-omniscient presence within the domestic sphere. As one popular 1861 guide to household management stated, '[a woman's] spirit will be seen through the whole establishment', suggesting the

---

[23] Millet, 'The Debate Over Women', 136.

[24] Coventry Patmore, *The Angel in the House* (London: Cassel & Company, 1891), 39, 38, 32, 38, Project Gutenberg.

[25] Dickerson, *Victorian Ghosts*, 6.

[26] Ibid, 4–8.

[27] Christina Rossetti, 'After Death', in *Goblin Market and Other Poems* (London: Macmillan and Co, 1865), 59.

need for a wife to be figuratively bodiless in order for her to have a thorough command over her household.[28]

Her and *BR2049* translate the ostensibly diminished corporeality of the Victorian wife into a literal spectrality. Wholly comprised of digital software, Samantha and Joi themselves resemble ghosts in their inability to interact with and manipulate the physical world. In her highly influential theory regarding the female body in cinema, Laura Mulvey argued that women assume an 'exhibitionist role' in mainstream films, in that the female body is positioned so as to be 'simultaneously looked at and displayed'.[29] In the context of this theory, it is interesting to note how both Jonze and Villeneuve use cinematic techniques to highlight Samantha and Joi's *lack* of solid body.[30] Joi, for example, is first introduced into the narrative of *BR2049* via a disembodied, off-screen voice; we hear her call out to K upon his return from work, 'I didn't hear you come in. You're early.' The camera proceeds to follow K around the claustrophobic spaces of his apartment as he continues to converse with the off-screen voice. As film theorist Mary Ann Doane notes, the efficacy of the off-screen voice in narrative cinema 'rests on the knowledge that the character can easily be made visible by a slight reframing which would reunite the voice and its source.'[31] Villeneuve manipulates this tension between voice and source by presenting a series of 'reframings' within K's apartment which do not reunite the feminine voice with a body, thereby creating an uncanny sense of a liminal, pervasive presence. When Joi's 'body' does eventually appear, it flickers out of the ceiling projector into the middle of the frame, conveying Joi as a fundamentally bodiless being, her holographic form a mere façade.

The use of the off-screen voice to create a sense of spectrality is even more pronounced in Jonze's *Her*. Like Joi, Samantha is first introduced into the film via a disembodied voice, in a private, domestic setting. In Theodore's first encounter with Samantha, after he loads the OS1 program into his computer, Jonze deliberately eschews the conventional shot-reverse-shot rhythm used in the earlier dialogue between Theodore and his blind date. Instead, the camera remains on a mid-shot of Theodore throughout the conversation; when Samantha is speaking, the camera does not reverse to the computer and thereby locate her 'being' inside of it. Although she tells the four-year-old Jocelyn that she 'live[s] in a computer', Samantha does not interact with or exist within the material realm at all; as Donna Kornhaber puts it, Samantha 'inhabits systems of data architecture and is contained within

---

[28] Isabella Beeton, *The Book of Household Management* (London: Jonathan Cape Limited, 1861), 1.

[29] Laura Mulvey, 'Visual Pleasure and Narrative Cinema', *Screen* 16, no. 3 (1975): 9.

[30] The implications of Mulvey's theory in regards to the filmic representation of digitised, bodiless women is discussed by Andrea Virginás in her article 'Gendered Transmediation of the Digital from *S1m0ne* to *Ex Machina*: "Visual Pleasure" Reloaded?', in which she argues for the 'ongoing validity of Laura Mulvey's concept of "visual pleasure" in spite of the apparently highly progressive medial re-codings [of female characters] in a digital environment'.

[31] Mary Ann Doane, 'The Voice in the Cinema: The Articulation of Body and Space' in *Narrative, Apparatus, Ideology: A Film Theory Reader* ed. Philip Rosen, (New York, Columbia University Press, 1986), 340.

them the way, we might say, a fish inhabits and is contained within the sea; which does not make those systems themselves her body'.[32] Samantha's ghostliness is reinforced in the revelation that she has been conversing with Alan Watts, a deceased philosopher whose consciousness was replicated as an artificial intelligence by group of OSes.

Samantha's ability to raise and speak with the dead further aligns her with the uncanny and the supernatural, suggesting that, like the heroine of Rossetti's poem, she occupies an ambiguous, liminal position between the material and spiritual worlds. Samantha and Joi's ghostliness and incorporeality parallels the limited material power of the Victorian woman removed by male rhetoric and technological development from the public and productive sphere. Moreover, the dispersed consciousness of these digital women reflects Victorian conceptions of the ideal wife as exuding a pervasive, ameliorative influence within the domestic space.

A reformative moral influence

The corporeality of the wife in Victorian Britain was de-emphasised as part of a growing perception of women as the stabilising moral centre of the home. In her natural propensity for emotion and feeling, it was seen as the woman's duty to render the hearth a sacred, calm space, and thereby civilise and ameliorate the emotional state of her husband. The idea that women performed a vital moral duty for men was propounded by John Ruskin, who even suggested that women were 'answerable' for past warfare, 'not in that [they] have provoked, but in that [they] have not hindered' conflict between men.[33] The reformative capacity of women was also reflected in literature: in 'The Angel in the House', Patmore likens his wife to an angelic visitor with transformative, almost godly capabilities; he writes that he is 'by her gentleness made great', and that 'she seem'd expressly sent below/To teach our erring minds to see'.[34] The reformative powers of the Angel in Patmore's poem, as well as broader conceptions of women's moral superiority in Victorian society, were inextricably connected to her diminished corporeality. Symbolising her personal physical and sexual desires, as well as her capacity for activity instead of passivity, a woman's active body was a disruptive incursion within the sacred domestic space. Conceptions of the ideal Victorian woman as the natural 'stabiliser of the home' therefore required that she '[disappear] into the woodwork … by becoming a ghost.'[35] In addition to the suppression of her physical form, however, her role as stabilising centre of the home also required the suppression of her personal desires and self-development.

---

[32] Donna Kornhaber, 'From Posthuman to Postcinema: Crises of Subjecthood and Representation in Her'. *Cinema Journal* 56, no. 4 (2017): 7.

[33] Ruskin, quoted in Millet, 'The Debate Over Women', 138.

[34] Patmore, *The Angel in the House*, 25, 21.

[35] Dickerson, *Victorian Ghosts*, 5.

Samantha and Joi both exemplify the idealised self-denial of the Victorian 'Angel' in that their *raison d'être* is the assistance and reformation of their respective male partners. As suggested in the epithet 'Angel' itself, the ideal Victorian woman was innately self-denying, perpetually subordinating her personal desires to the needs of her husband and family. It was a widely held conception in Victorian society that while men worked to satisfy both public and self-interest, women's labour was always for others; as Elaine Showalter notes, 'work, in the sense of self-development, was in direct conflict with the subordination and repression inherent in the feminine ideal'.[36] Joi, in particular, is shown to possess little individual subjectivity outside of her relationship to K. Silently and invisibly contained within the emanator, Joi's existence is wholly tied to and reliant upon K's being, as reflected in her rooftop profession: 'I'm so happy when I'm with you.' While *Her*'s Samantha exhibits a greater degree of self-awareness and independence than Joi, her relationship with Theodore is nevertheless considerably unbalanced owing to the inherent incongruities in their bodily and cognitive forms. When Samantha tells Theodore that she is 'gonna be really lonely when [he] sleep[s],' she, like Joi, presents her own personhood as dependent on his presence. Samantha and Joi thus reflect the idealised Victorian image of the self-sacrificing woman in that their primary function and purpose lies in ancillary work for their 'husbands'.

In their programmed self-denial, both Joi and Samantha devote themselves to providing emotional stability and support to their respective male partners, and fostering the men's personal and social development. Theodore in particular is presented as wholly 'underdeveloped' prior to his purchase of Samantha – he is likened to 'a little puppy dog' by his blind date, and seems to struggle with forming romantic connections following his divorce.[37] The OS1 program therefore appeals to Theodore in its promise of personal development and regeneration: the advertisement in the train station rhetorically asks, 'Who are you? What can you be? Where are you going? … What are the possibilities?' Theodore's purchase of Samantha offers him the regeneration he has been seeking; she not only stabilises his life through her secretarial assistance, but also reforms his character on a deeper level by providing him an opportunity to vocalise and work through his internal emotional conflict. As Matt Aibel argues, Samantha's 'mirroring and validation … gradually help [Theodore] consolidate his previously isolated or fragmented self states'.[38]

Samantha's invisible, salutary presence also continually recreates for Theodore the comforts of the domestic space. Like the 'Angel' of Patmore's poem, Samantha's reformative influence is intimately connected to her lack of body: a constant external conscience, she watches over Theodore while he

---

[36] Elaine Showalter, *A Literature of Their Own: British Women Novelists from Bronte to Lessing*. (London: Virago, 1977), 22.

[37] Matt Aibel, 'From Provisioning to Reciprocity: Logging in to Spike Jonze's *Her*'. *Psychoanalytic Psychology* 34, no. 3 (2017): 369.

[38] Ibid, 369.

sleeps, accompanies him to all his social outings, and is available to talk with him whenever he desires. John Ruskin's 1864 declaration that 'wherever a true wife comes, [the] home is always round her' is thus exemplified in Samantha, who in her constant, bodiless companionship of Theodore exudes the comforting energy of a stable, sacred hearth.[39]

Joi similarly embodies the reformative powers of the Victorian 'Angel' in that her spectral presence facilitates K's emerging sense of individual selfhood. Like the perpetual companionship Samantha offers Theodore, K is able to summon Joi into existence any time he requires guidance or support. She appears, for example, as he tries to interpret the biological database, and as he sets out on his perilous flight to the orphanage. Even more so than Theodore, K possesses an underdeveloped sense of individual subjectivity and selfhood. A replicant officer in the Los Angeles Police Department, K is referred to by the moniker 'Officer K6-3.7', and it is apparent from his first entry into the narrative that he has internalised the soulless spirit of mass production embodied in such a generic name. Like the ideal, supportive wife in the emerging industrial market of the Victorian era, Joi encourages K to see himself as extraordinary; she tells him he is 'special' and 'a real boy', and, most significantly, offers him a new name, Joe. As the French phenomenological philosopher Maurice Merleau-Ponty argued, to name someone is to 'accredit objectivity, self-identity, positivity, [and] plenitude'.[40] As such, when Joi tells K, 'you're too important for K. Your mother would have named you Joe', it indicates her crucial role as a reformative and constructive moral influence.

In naming K herself, Joi assumes the mantle of the absent mother, reflecting the dual role of the Angel in the House as both wife and mother, tasked with the moral education of her children as well as the amelioration of her husband's emotional state. Like Samantha, then, Joi provides K with a 'crucially needed self-object experience' through her constant reformative presence.[41] Overall, the digital women of *Her* and *BR2049* reflect an idealised image of womanhood which locates female purpose in the provision of ancillary, reformative domestic support to men. Furthermore, in their programmed self-denial and their contained omnipresence, Joi and Samantha appeal to an idealised perception of women as static and therefore stabilising presences within the domestic sphere, especially amidst rapid and unstable technological change in the public sphere.

Conclusion

Comparing the Victorian image of the Angel in the House with the digital women in *Her* and *Blade Runner 2049* suggests an iterated pattern in constructions of ideal womanhood associated with

[39] Ruskin quoted in Millet, 'The Debate Over Women', 131.

[40] Maurice Merleau-Ponty, *The Visible and the Invisible,* trans. A Lingus (Evanston: Northwest University Press, 1968), 162.

[41] Aibel, 'Logging in to *Her*', 369.

technological development. In both the Victorian era and the futuristic societies of these speculative fiction films, rapid technological and social development, driven by men, has destabilised the female subject position. The ideal woman which emerges in both Victorian Britain and the societies of *Her* and *BR2049* is contained with patriarchal structures of power and is removed from the productive and material world. This ideal woman also possesses an ambiguous and insubstantial corporeality, which allows her to imbue the domestic sphere with a pervasive feminine influence. As part of this influence, she subordinates her own self-development and desires to a higher purpose, namely, the reformation of her male partner. The parallels between these two spatially and temporally distant constructions of ideal womanhood discourage conceptions of a linear progression of gender relations or gendered subject positions. Rather, they speak to a broader iterative pattern of female experience that exists regardless of time and space.

Acknowledgements

I would like to thank Dr Kate Flaherty for her willingness to supervise me in this research project and for her guidance in writing my essay.

Bibliography

Aibel, Matt. 'From Provisioning to Reciprocity: Logging in to Spike Jonze's *Her*'. *Psychoanalytic Psychology* 34, no. 3 (2017): 368–71. doi.org/10.1037/pap0000121.

Beeton, Isabella. *The Book of Household Management.* London: Jonathan Cape Limited, 1861.

Berg, Maxine. 'What Difference did Women's Work Make to the Industrial Revolution?' *History Workshop* 35 (1993): 22–44. doi.org/10.1093/hwj/35.1.22.

Bordun, Troy. 'On the Off-Screen Voice: Sound and Vision in Spike Jonze's *Her*'. *Cineaction* 98 (2016): 57–64.

Cockburn, Cynthia. *Machinery of Dominance: Women, Men and Technical Know-How*. London: Pluto Press, 1985.

Cordea, Diana. 'Two Approaches on the Philosophy of Separate Spheres in Mid-Victorian England: John Ruskin and John Stuart Mill'. *Procedia: Social and Behavioural Sciences* 71 (2013): 115–22. doi.org/10.1016/j.sbspro.2013.01.016.

Dickerson, Vanessa. *Victorian Ghosts in the Noontide.* Columbia: University of Missouri Press, 1996.

Doane, Mary Ann. 'The Voice in the Cinema: Articulation of Body and Space'. In *Narrative, Apparatus, Ideology: A Film Theory Reader,* edited by Philip Rosen, 162–76. New York: Columbia University Press, 1986.

Griffin, Ben. 'Class, Gender, And Liberalism in Parliament, 1868–1882: The Case of the Married Women's Property Acts'. *Historical Review* 46, no. 1 (2003): 59–87.

Jollimore, Troy. 'The Endless Space Between the Worlds: The Limits of Love in Spike Jonze's *Her*'. *Midwest Studies in Philosophy* 39 (2015): 120–43. doi.org/10.1111/misp.12039.

Jonze, Spike, dir. *Her*. Culver City: Sony Pictures Home Entertainment, 2013. DVD.

Kirkup, Gill. 'Introduction'. In *The Gendered Cyborg: A Reader*, edited by Gill Kirkup, Linda James et al, 3–10. London: Routledge, 2000.

Kornhaber, Donna. 'From Posthuman to Postcinema: Crises of Subjecthood and Representation in *Her*'. *Cinema Journal* 56, no. 4 (2017): 3–25. doi.org/10.1353/cj.2017.0038.

Merleau-Ponty, Maurice. *The Visible and the Invisible*, translated by A Lingus. Evanston: Northwest University Press, 1968.

Millet, Kate. 'The Debate Over Women: Ruskin vs Mill'. In *Suffer and Be Still: Women in the Victorian Age,* edited by Martha Vicinus, 121–39. Bloomington: Indiana University Press, 1973.

Mulvey, Laura. 'Visual Pleasure and Narrative Cinema'. *Screen* 16, no. 3 (1975): 6–18.

Patmore, Coventry. The Angel in the House. Urbana: Project Gutenberg, 2014. http://www.gutenberg.org/files/4099/4099-h/4099-h.htm.

Perkin, Joan. *Victorian Women.* London: John Murray, 1993.

Peterson, Mildred Jeanne. *Family, Love, and Work in the Lives of Victorian Gentlewomen.* Bloomington: Indiana University Press, 1989.

Rossetti, Christina. 'After Death', in *Goblin Market and Other Poems*, 59. London: Macmillan and Co, 1865.

Schreiner, Olive. *Women and Labour.* Mineola: Dover Publications, 1998.

Showalter, Elaine. *A Literature of Their Own: British Women Novelists from Bronte to Lessing*. London: Virago, 1977.

Vicinus, Martha. 'The Perfect Victorian Lady'. In *Suffer and Be Still: Women in the Victorian Age*, edited by Martha Vicinus, vii–xv. Bloomington: Indiana University Press, 1973.

Villeneuve, Denis, dir. *Blade Runner 2049*. Culver City: Sony Pictures Home Entertainment, 2017. DVD.

Virginas, Andrea. 'Gendered Transmediation of the Digital from *S1m0ne* to *Ex Machina*: "Visual Pleasure" Reloaded?' *European Journal of English Studies* 21, no. 3 (2017): 288–303. doi.org/10.1080/13825577.2017.1369266.

Wrigley, Antony. *Continuity, Chance and Change: The Character of the Industrial Revolution in England*. Cambridge: Cambridge University Press, 1990.

Dear, Aphro[die]te

DANA ROYLE

Abstract

Complexity encompasses the entirety of social life, which inspired me to utilise artistic forms, specifically sculpture, as a way to explore these ideas. The piece that I created (Figure 1), entitled *Dear, Aphro[die]te*, is an attempt to unmask our own ignorance surrounding different kinds of privilege, focusing on the enduring nature of gender inequality and the rights of women. This is not a subject that one can fully understand or solve while holding a singular perspective, and so to inspire more analytical thought on this subject, I created a modern tale describing the demise of Aphrodite. In this sense, the complex issue of gender inequality can be experienced in a way that one may not have felt before, allowing for the exploration of different perspectives and patterns of thought, observing the subject with new eyes and creating innovative ways to address gender inequities.

Introduction

Aphrodite was once the Ancient Greek goddess of beauty, pleasure, desire, love, and sexuality. She was once an immortal being, holding great power in both her looks and her ability to make others fall in love. However, due to Aphrodite's persistent meddling in the affairs of mortals, she was banished by Zeus to live among them as a test of her identity. Now facing the reality of death, Aphrodite looked towards the mortals for help as to save her from this detrimental confrontation, and to reclaim a grip on her status. She soon found that she was no match for the values, pressures, and expectations of this new world she was trapped in. Even though Aphrodite had already lost grip of her true potential, she became further isolated and suppressed, stripping her of any remaining fragment that once made up her essence. She was not valuable. She was not important. She was not to be heard. She was worthless. She was unknown. She was broken.

I created a sculpture (Figure 1) to exhibit the ultimate ending of this modern tale of Aphrodite when she became situated in contemporary society.

Figure 1: Dana Royle, *Dear, Aphro[die]te*. 2018. Mixed media.

Source: Photograph by Dana Royle.

I created this sculpture as a visual critique of modern society, in terms of the complex nature of gender inequality and the experiences of women in everyday life. It features two of the main aspects of the story of Aphrodite. First is the moment of Aphrodite's banishment by Zeus to live among mortals, where she is completely stripped of the foundations of her identity, leaving her vulnerable and searching for help. Second is her ultimate and unremarkable demise in the corrupt world she finds herself trapped in. To illustrate this, this paper begins by discussing how the individual aspects of sculpture are representative of Aphrodite's experience in modern society. I then go into the wider issue of gender inequality and its complex and enduring nature through all aspects of life. Finally, I explore how art can be utilised as a medium to address and ultimately solve complex problems such as that of gender inequality. I arrive at the conclusion that it is through the expression of complex problems in art, pertaining to *Dear, Aphro[die]te,* that different perspectives can be explored, observing the subject from a new perspective and forming innovative ways to address complex issues.

The face of Aphrodite was purposefully designed to gaze back at the viewer who is gazing at her, as to be one that confronts their beholder, establishing the power and authority of the audience over the artwork, and more specifically over Aphrodite (Olin, 1996). She looks outwards to the audience for help, with her last desperate hope, as she is left powerless and at the mercy of the audience. Even though all the power is given to the audience to try and make a change, they are incapable of anything other than feeling confronted by their own sense of superiority.

The blandness in detailing and colour of Aphrodite's face is further representative of her loss of identity and soul, allowing the audience to gaze at her dehumanised face guilt-free, as she cannot see what the audience sees. She cannot see that the audience has failed her, and watched her demise. This gaze of the audience itself illustrates the public's part to play in the unavoidable future of Aphrodite, represented by the graveyard scene on top of Aphrodite's head. The audience in this sense is the graveyard worker, bringing Aphrodite's life to a close, with their thoughts only within themselves, ready to move on and take no real notice or responsibility for what has happened.

The sculpture pushes gender representations to an extreme and takes a note from feminist art by examining the female identity through a personalised narrative (Fineberg, 1995). The aim is to bring to the forefront the suppressed perspective of the female within the patriarchal society that we have today, exposing these experiences to the audience for their consideration and awareness. In this sense, it serves as an evocative form of social protest, challenging the institutionalised value systems and social constructions of what it means to be a woman (Broude & Garrard, 1994). The conventional image and representation of women is one that is completely passive in nature; to be seen but not heard (Fineberg, 1995). She should exist only as an object of possession for masculine gaze, fantasies and desires (Olin, 1995). She must appear beautiful without trying, her looks being set to an unattainable standard, with any attempt to try and meet this ideal being futile (Brand, 1998).

It is these concepts that this sculpture aims to represent; the repression and subjugation of any kind of female empowerment. It works to challenge and deconstruct these preestablished conventions of womanhood and femininity, for just as Aphrodite lost her true identity as the goddess of love and ultimately became isolated, women are weakened by the internalised demands of a male-dominated society (Nochlin, 1989). It illustrates how women are tied to these implications of the patriarchy, confronting its audience with the reality of the world we live in, and serving as a forewarning of society's mistreatment of women, aiming to encourage and reinforce the need for change in the existing ignorant culture.

This discourse on the value of women is just the tip of the iceberg in terms of the complex nature of gender inequality to which *Dear, Aphro[die]te* speaks. The simple difference in one chromosome leads to women having a one in three chance of experiencing sexual or domestic violence (Australian Bureau of Statistics, 2016) and to be paid on average 15 per cent less than their male counterparts (Australian Bureau of Statistics, 2018). More widely, the UN Department of Public Information (2011: 1) further reported that in the world, 'women aged 15–44 are more at risk from rape and domestic violence than from cancer, car accidents, war and malaria'. Even within the art sphere itself, a space inherently labelled as feminine, these gender boundaries can be seen, with women's voices discouraged and underrepresented. For example, in a book published in 1962 titled *History of Art*, the authors failed to name a single female artist nor were any female artists' works exhibited (Broude & Garrard, 1994). This underrepresentation of women expresses the idea of the 'inability of human beings with wombs rather than penises to create anything significant' (Nochlin, 1989: 146). This makes it clear that the question of gender inequality cannot simply be limited to or explained by

> the relative benevolence or ill-will of individual men, nor the self-confidence or abjectness of individual women, but rather on the very nature of our institutional structures themselves and the view of reality which they impose on the human beings who are part of them. (Nochlin, 1989: 152)

It is through the use of art as a medium to address and explore the complex subject of gender inequality that new, innovative ways of awareness can be created (Kelly, 1996). Just like complexity itself, the format and presentation of art is subjective and hard to define, allowing differing meanings, interpretations, experiences, knowledges and understandings to be ascribed to a single artistic piece (Broude & Garrard, 1994). By subjecting complex issues to this format, it is possible to obtain the attention of the multidisciplinary perspectives which are required to address the issue, start a conversation, and further begin to work together towards a solution. Art has the unique ability to yield multiple, and often conflicting, perspectives, forcing these ideas to be triangulated together, making the audience look further for the solutions to complex problems, toward our inner selves, predispositions, and institutional structures (Brand, 1998). Art can also act as a lens of one's culture, institutions and education, allowing the artist to create a piece that serves as an effective and evocative form of social protest.

It is through the visualisation in art pertaining to *Dear, Aphro[die]te* that it is possible to question what has been accepted as the 'normal', and to challenge the systems that are currently in place. This may transform cultures in widespread and meaningful ways (Nochlin, 1989). For just as complex problems require complex solutions, finding them also needs complex representations and complex understandings. Therefore, art can and should be utilised as a medium to explore possible solutions to complex issues through its aim of representation. Not only will this make ideas and concepts more accessible, allowing complex issues to be communicated more widely, but it will create a combined understanding in its complex nature. By expressing complex issues through art, it is possible to inspire deeper thought and evaluation about change, improvement, and progress, serving as developments towards the avant-garde movement (literally translated as 'to the forefront').

Acknowledgement

Thank you to my lecturer, Chris Browne, for the freedom and inspiration to create this piece. Just because something is complex, it does not mean it can be ignored.

References

Australian Bureau of Statistics. (2016). *Personal Safety Australia: Key Findings*. (cat. no. 4906.0), www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4906.0~2016~Main%20Features~Key%20Findings~1.

Australian Bureau of Statistics. (2018). *Average Weekly Earnings: Nov 2017* (cat. no. 6302.0), www.abs.gov.au/ausstats/abs@.nsf/mf/6302.0.

Brand, P, (1998). 'Disinterestedness and Political Art'. In *Aesthetics: The Big Questions* (ed. Carolyn Korsmeyer), Blackwell publishers, Massachusetts.

Broude, N, and Garrard, M. (1994). *The Power of Feminist Art: The American Movement of the 1970s*. Harry N Abrams, New York.

Fineberg, J. (1995). *Art Since 1940: Strategies of Being*. Laurence King Publishing, London.

Kelly, M. (1996). *Imaging Desire*. MIT Press, Cambridge Massachusetts.

Nochlin, L. (1989). 'Why Have There Been No Great Women Artists?' In *Women, Art and Power and Other Essays* (ed. Linda Nochlin), Thames and Hudson, London.

Olin, M. (1996). 'Gaze'. In *Critical Terms for Art History* (ed. Roberts Nelson and Richard Shift), University of Chicago Press, Chicago.

UN (United Nations) Department of Public Information. (2011). *United Nations Secretary-General's Campaign: Unite to End Violence Against Women*. DPI/2546A, 16dayscwgl.rutgers.edu/downloads/2011-campaign-docs/198-2011unite-pdf/file.