How well do psychologists' research methods equip them to identify the impacts of climate change on behaviour? A methodological investigation with particular reference to the effects of temperature on violent behaviour

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy
in
Psychology

at Massey University, Albany,
New Zealand

Matthew Neil Williams

2015

Abstract

The Earth's temperature is rising, and it is extremely likely that human activities are primarily to blame (IPCC, 2013b). A changing climate could have serious consequences for human behaviour and psychological functioning. Research concerned with the psychological impacts of climate change is challenging, however, given the paucity of data showing how human behaviour has responded to sustained climate changes in the past. In this thesis, I critique the suitability of psychologists' mainstream methodological strategies for engaging in research concerned with the impacts of climate change. In doing so, I draw heavily on a specific "testbed" of psychological research relevant to climate change: Research concerned with the impact of temperature on the incidence of intra- and interpersonal violence. In identifying methodological problems I draw both on published literature as well as an empirical engagement in research in this area. The empirical component constitutes an analysis of the relationship between temperature and the incidence of assault, suicide, and self-harm resulting in hospitalisation in New Zealand. In this analysis I found that irregular day-to-day variation in temperature had a positive relationship with all three forms of violence. However, there was less evidence that more sustained (seasonal or geographical) differences in temperature led to increased violence, making it difficult to predict the effects of sustained increases in temperature in the future. In the methodological critique section of this thesis, I point out several methodological problems that may hamper psychologists' capacity to produce effective and useful research concerned with the impacts of climate change. These problems include the use of measurement and analysis strategies that limit our ability to convey the sizes of effects; the use of theories and analyses that limit our ability to make predictions; and the inadequate reporting of uncertainty. Finally, I recommend that psychologists studying climate change impacts should consider using categorisations of behaviour rather than psychometric scales that lack clear units of measurement; use statistics that effectively communicate effect size; apply theories that facilitate prediction-making; carefully take into account the role of time when generating predictions; and account for multiple sources of uncertainty that affect the confidence of our conclusions.

Acknowledgements

This thesis may have my name on it, but many people have contributed to it. First thanks go to my supervisors Stephen Hill and John Spicer, who have been patient and helpful sources of wisdom throughout. I couldn't have asked for two better supervisors, and I certainly hope this isn't the last time we work together. I'd also like to thank the editors and reviewers who provided helpful feedback on the journal articles presented as part of this thesis. Any errors that remain are my own. Earlier in the process of my PhD, several organisations were very helpful in providing me with data, including the New Zealand Police, the Ministry of Health, the National Institute of Water and Atmospheric Research (NIWA), and Statistics New Zealand. Of particular help were Lynn Jenner, Gavin Knight, Steve Darroch, and Ruth Wallis (NZ Police); Jane Perrott and Chris Lewis (Ministry of Health); and Richard Speirs (Statistics NZ).

Margaret Roberts, Jaimie Veale, and Kirsty Furness were of great help when proofreading this thesis. Alison McKinlay deserves a lot thanks for sharing an office with me over these last few stressful months. Even more importantly, having the company of you four (and many others) also going through the trials and tribulations of doctorate study helped a lot. I'd also like to thank the inhabitants of the talkstats.com forums. What little I know about statistics is mostly thanks to all of you.

My parents Anneke and David have been amazing sources of love and support (and are great proofreaders too!) Thanks also to my brother and sister, Martin and Amy, for always being excellent and hilarious company. While writing of family, I'd also like to acknowledge my uncle Herman Gerhardus van Gysen, a researcher much more accomplished than I am who sadly died at the age of just 45. Herman's research was mainly focused on spline models in geodesy. I used spline models briefly in this thesis (see section 3.4.1), and I'm sure my analyses would have been much improved if I'd had him around to talk with.

Some PhD students manage to focus solely on their research for years at a time; I am not one of those students. I'd like to thank the volleyball crew for always being there when I needed a break and some exercise. The same goes for the longtime friends like Kris, Nick, and Rob who've been there since I first decided I'd like to spend another year… or two… at university.

Last but not least, I'd like to thank my wonderful partner Jessica. Finally, it's time for us to have a holiday.

Preface

This thesis takes the form of a partial thesis-by-publications format. Specifically, the empirical section of the thesis is presented as three journal articles. Studies One and Two were both published (in print and online) in *Climatic Change* in 2015. Study Three was published in *Psychology, Health, & Medicine* as an advance online publication in 2015. The final two chapters of this thesis (the methodological critique and methodological recommendations chapters) have not been submitted for publication. I chose not to structure these chapters as manuscripts in order to take advantage of the extra freedom in terms of space and format that a thesis allows. An additional short commentary article that I published in 2013 along with my supervisors, but that does not form a part of the main narrative of this thesis, is included in Appendix D.

The publishers of the three main articles presented in this thesis (Springer for *Climatic Change* and Taylor & Francis for *Psychology, Health & Medicine*) both provide authors with the right to include published articles in a thesis or dissertation. I contacted both publishers to confirm that this was acceptable in my specific case (i.e., a thesis that will ultimately be accessible online, with minor formatting changes to the articles presented). Representatives of both Springer and Taylor & Francis kindly confirmed that this was acceptable. The commentary article presented in Appendix D was published in an open access journal (the *Western Journal of Emergency Medicine*), meaning that it could be reproduced without obtaining permission from the publisher.

The work presented in this thesis is my own. I designed the empirical studies, collected data, selected and conducted data analyses, and wrote all of the chapters presented. My supervisors Stephen Hill and John Spicer helped me to select an appropriate structure, provided feedback on my writing, and provided valuable advice with respect to conceptual issues. They were therefore included as co-authors for the publications included in this thesis.

Ethical approval for the empirical studies reported in this thesis was obtained from the Massey University Human Ethics Committee, Southern B Application 10/54.

Matt Williams
Doctoral Candidate
Massey University

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA/ANCOVA	Analysis of variance/Analysis of covariance
AWS	Automatic Weather Station
CI	Confidence interval
df	Degrees of freedom
ENSO	El Niño Southern Oscillation
ICD	International Classification of Diseases
IPCC	Intergovernmental Panel on Climate Change
M	Mean
MJ/m^2	Megajoules per square metre (of radiation)
NASA	National Aeronautics and Space Administration
NIWA	National Institute of Water and Atmospheric Research
NZ	New Zealand
RCP	Representative Concentration Pathway
SD	Standard deviation
SI	International System of Units
SOI	Southern Oscillation Index
Tmean	Mean temperature
US/USA	United States of America
VCN	Virtual Climate Network

1 Introduction: Climate Change and Psychology

> An urgent geophysical fact has become clear. Burning all the fossil fuels will destroy the planet we know, Creation, the planet of stable climate in which civilization developed.
>
> Hansen and Hansen (2008)[1]

Modern humans originated in Africa around 200,000 years ago (Lewin, 1987). Remarkably, most of the major features of current human civilisation—including agriculture, cities, and written language—were all developed in the Holocene, a period of climatic stability that began approximately 11,700 years ago (see Dansgaard et al., 1993; Gupta, 2004; Leick, 2002; Li, Harbottle, Zhang, & Wang, 2003). That climatic stability is now under serious threat. The fifth report of the Intergovernmental Panel on Climate Change (IPCC, 2014a) found that evidence for warming of the global climate is "unequivocal" (p. 1), and that it is "extremely likely" (p. 4) that human activities are the dominant cause of this warming.

While anthropogenic greenhouse gas emissions are of course likely to result in rising temperatures (IPCC, 2013b), a phenomenon known as "global warming", they also have a number of other consequences. The term "anthropogenic climate change" refers more broadly to the wide range of climatic changes that are currently occurring, or that are likely to occur in the future, as a result of human activities. These climatic changes involve serious environmental risks. Likely effects over the next century include rising temperatures and sea levels, an increased rate of species extinctions, and an increased risk of floods (IPCC, 2014a).

Given that psychology is the study of mind and behaviour (American Psychological Association, 2012), psychologists might be expected to make a significant contribution to advancing understanding of how human behaviour may be affected by climate change, and how human behaviour might be altered to mitigate the threat of climate change. Yet psychology's engagement in research concerned with climate change has been relatively limited to date (Gifford, 2008). Furthermore, the

[1] This quote is from a letter by climatologist James Hansen, co-signed by Dr. Hansen's wife Anniek, and addressed to Barack and Michelle Obama.

capacity of psychologists to identify and quantify the effects of climate change on human behaviour and psychological functioning is unclear. The extent and success of psychologists' engagement with the problem of climate change will depend on a variety of factors, not least psychologists' capacity to deal with the significant methodological difficulties involved in climate change research. As a multifaceted, extensive, and potentially severe change to the human environment, climate change presents a number of specific constraints and drivers that make its study a particular challenge for psychologists.

This introductory chapter will begin with a brief review of the evidence for anthropogenic climate change, and how climate change is likely to affect humans' physical health and habitat (section 1.1). This will be followed by an argument for why psychologists need to engage in climate change research, including study of the impacts of climate change on human behaviour and psychological functioning (section 1.2). Finally, in sections 1.4 and 1.5, I will outline the methodological challenges that studying climate change impacts may present for psychologists, and explain how this thesis will examine psychology's readiness to deal with these challenges.

1.1 Anthropogenic Climate Change: Consensus and Consequences

1.1.1 The evidence for anthropogenic climate change.

The hypothesis that increased concentrations of greenhouse gases might eventually lead to warming of the Earth's atmosphere is not a new one. The greenhouse effect—the process wherein gases such as carbon dioxide and methane trap heat in an atmosphere by absorbing and re-radiating thermal radiation from a planetary surface—was first proposed by Joseph Fourier in 1827, and demonstrated experimentally by John Tyndall in 1859 (Held & Soden, 2000). Before the turn of the 20th century, the Swedish scientist Svante Arrenhius (1896) suggested that increases in the carbon dioxide concentration in the atmosphere might lead to warming of the global climate. Arrhenius (1908) later went on to suggest that humans' widespread burning of fossil fuels might have just this effect.

More than half a century later, observations at the Mauna Loa observatory in Hawaii provided the first strong evidence that levels of carbon dioxide in the Earth's atmosphere were indeed rising (Keeling, 1960). This observation resulted in gradually increasing concern about the climatic implications of a change to the composition of the

Earth's atmosphere. In 1988, the Intergovernmental Panel for Climate Change (IPCC) was created by the United Nations to investigate the risk of anthropogenic climate change and its potential consequences. As evidence has accumulated, the IPCC's reports have expressed more and more confidence that the Earth is warming, and that human activities are responsible for this warming. As mentioned above, the most recent (fifth) IPCC report concluded that evidence for global warming is "unequivocal" (p. 1), and that it is "extremely likely" (p. 4) that human activities are the dominant cause of this warming (IPCC, 2014a). The IPCC define the term *extremely likely* as indicating a probability of greater than 95% (Mastrandrea et al., 2011). The recent warming of the Earth's temperature is clear in Figure 1.

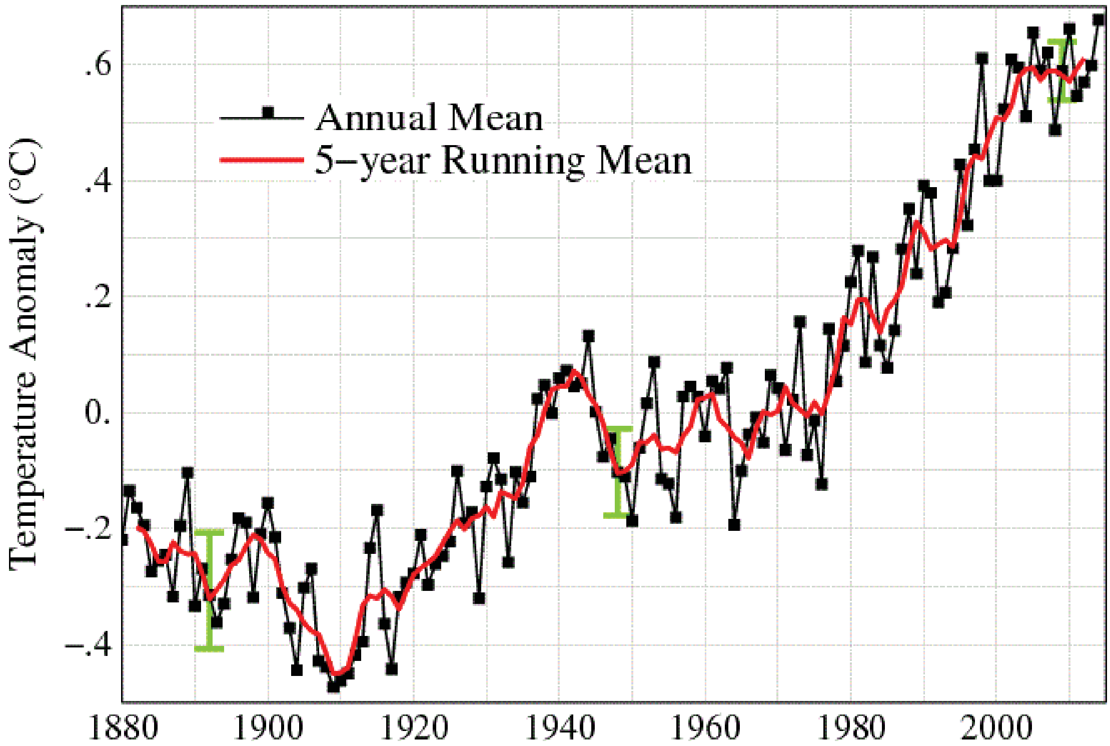

Figure 1. Changes in the global surface air temperature. The dotted black line shows the annual mean land-ocean temperature index. The red line shows the five-year running mean and the green bars represent uncertainty estimates. Temperature anomalies are relative to a 1950–1980 base period. The figure is an updated version of that found in Hansen, Ruedy, Sato and Lo (2010), and obtained from the National Aeronautics and Space Administration (NASA; n.d.). Image in the public domain.

The IPCC's position that the Earth is warming and human activities are responsible for most of this warming is backed by a very strong scientific consensus. An extensive review of the publications of 1372 climate scientists found that 97–98% of these scientists supported this conclusion (Anderegg, Prall, Harold, & Schneider, 2010). Similarly, a review of 11,944 abstracts published in scientific journals between 1991 and 2011 matching the search terms "global climate change" or "global warming" found that of the papers that expressed a position about anthropogenic global warming, 97.1% agreed that humans are causing warming of the global climate (Cook et al., 2013).

1.1.2 Future trajectory and consequences of climate change.

The increase in temperature likely over the following decades and centuries depends heavily on human activities. One way that these influences can be taken into account in climate models is by considering a variety of scenarios for future greenhouse gas emissions. In the most recent IPCC report, this was accomplished using scenarios known as Representative Concentration Pathways (RCPs; Moss et al., 2008). For example, scenario RCP2.6 describes a world in which emissions peak this decade and decline substantially thereafter. Under this very optimistic scenario, temperature change by the period 2081–2100 (relative to the 1986–2005 mean) has a likely range of 0.3 to 1.7°C (IPCC, 2013b). On the other hand, the RCP8.5 scenario describes a world in which greenhouse gas emissions continue to rise throughout the 21st century. This scenario leads to a likely range for temperature increase over the 21st century of 2.6 to 4.8°C (IPCC, 2013b).

An increase in global temperatures is the most well-known component of anthropogenic climate change. However, the enhanced greenhouse effect occurring as a result of human activities has much wider consequences for the global climate than just increases in average temperature. Some of the changes projected to occur by the late 21st century in the most recent IPCC report include more heat waves, increases in the frequency and/or intensity of heavy precipitation events in most mid-latitude land areas, increases in the intensity and/or duration of droughts, and continuing ocean acidification (IPCC, 2013b).

1.1.2.1 Effects on human physical health.

The environmental changes occurring as part of climate change will in turn have implications for human health. These effects are likely to include mortality and morbidity due to an increased intensity of heat waves and fires; undernutrition due to reduced agricultural production in poor regions; and an increased risk of food-, vector- and water-borne diseases (K. R. Smith et al., 2014). Morbidity and mortality due to cold exposure may be reduced, but this effect is likely to be outweighed by the negative effects of increased heat stress; overall, while some of the effects of climate change on health will be positive, the net worldwide effect on health is expected to be negative (K. R. Smith et al., 2014). Some of these effects are being felt already: The World Health Organisation (2002) estimates that approximately 154,000 excess deaths worldwide could be attributed to climate change in the year 2000. These excess deaths occurred via effects of climate change on the incidence of illnesses such as diarrhoea, malaria, and dengue fever.

1.2 The Current and Potential Contribution of Psychology to Climate Change Research

Anthropogenic climate change is a problem that was first identified and quantified by researchers in the physical sciences, including physicists, meteorologists, geologists, and climatologists. These physical scientists have obviously played a very important role in climate change research, and continue to do so. However, there have recently been a number of calls for psychologists and social scientists to engage in climate change research (Gifford, 2008; Pidgeon & Fischhoff, 2011; Spence, Pidgeon, & Uzzell, 2009; P. C. Stern, 2011; Swim et al., 2011). One rationale for these calls is that the cause of anthropogenic climate change is a phenomenon that falls firmly in the domain of psychologists and other behavioural scientists: human behaviour. It is behaviour by humans that has resulted in the dramatic increase in atmospheric greenhouse gas concentrations since the pre-industrial period, and human behaviour in the future that is expected to lead to a continuation of global warming.

1.2.1 Psychological research and the mitigation of climate change.

One area in which psychologists can engage in climate change research is that of research concerned with the *mitigation* of climate change: that is, research concerned with how human behaviour can be changed to prevent or minimise climate change. Oftentimes, the findings of psychological research projects broadly concerned with the mitigation of climate change are bleak. Rachlinski (2000) argues that anthropogenic climate change represents a *commons dilemma* in that behaviours that support humans' individual, short-term interests (i.e., carbon emitting behaviours involved in transportation and economic production) are highly destructive for the species as a whole over the long term. A range of psychological denial mechanisms that help individuals to justify inaction on climate change were identified in a Swiss study (Stoll-Kleemann, O'Riordan, & Jaeger, 2001). Pawlik (1992) identified five major barriers to climate change action: the relatively small magnitude of mean temperature change compared to day-to-day variations; the large temporal delay between the actions that cause climate change and their eventual effects; our tendency to underestimate changes in the relative frequency of rare events such as heat waves and floods; the social distance between those who contribute most to causing climate change and those who will be most affected; and the conflict between environmental and economic imperatives in the short term.

Even where an individual believes that anthropogenic climate change is occurring, this may not necessarily translate into meaningful behaviour change to mitigate its effects. Gifford (2011) discusses several "dragons of inaction" (p. 290) that act as psychological barriers to prevent pro-environmental attitudes translating into actual behaviour change. These barriers include limited cognition about climate change, anti-environmental ideologies, behavioural inertia, and the perceived risks associated with behavioural change. Whitmarsh (2009) found that even where individuals do take actions intended to mitigate climate change, these actions may take the form of behaviour that is arguably pro-environmental but not specifically helpful for reducing greenhouse gas emissions (e.g., recycling).

1.2.2 The need for study of the psychological *effects* of climate change.

Research concerned with human behaviour and its effects on the climate is undeniably important. The knowledge produced by psychological research in this area may ultimately be of some help towards efforts to mitigate climate change. However, there is a need to be realistic: The Earth is already locked in to at least some warming over the next century, and it is likely that this warming will be substantial (see IPCC, 2013b). Certainly, psychological research concerned with climate change provides little reassurance that rapid behaviour change to mitigate climate change will happen in the immediate future. The likely consequences of climate change for human behaviour are worth investigating, particularly given the magnitude of the changes to humans' habitats that are involved.

Indeed, climate change may result in a dramatic and fundamental change to the environment that humans live in. Humans obviously have some ability to adapt to different climates, but within limited bounds. For example, global warming of more than 7°C could render some regions of the world uninhabitable, given that the temperatures in these regions would reach levels that exceed our ability to thermoregulate (Sherwood & Huber, 2010). Furthermore, despite the advances of modern infrastructure and health systems, humans—like most animals—are not well adapted to extreme weather events such as heat waves and floods. There appears to be a reasonable *a priori* basis to expect that environmental changes such as substantial temperature increases, sea level rise, and changes in the incidence of extreme weather events will have substantial effects on human behaviour and psychological functioning. Investigating what specific effects may be involved is useful so as to be able to provide information that could be used by policymakers and other stakeholders to help ameliorate any harmful psychological effects of climate change.

1.2.2.1 Pathways for psychological effects of climate change.

Climate change may affect human behaviour and psychological functioning in a number of different ways. Berry et al. (2010) proposed a model specifying three important causal pathways for such effects. Although this model focuses specifically on impacts on mental health, the pathways identified may very well result in effects on behaviour more broadly, as opposed to just mental health. As such, the model provides a useful framework for a discussion of how climate change may impact human behaviour and psychological functioning.

The first pathway via which climate change may affect mental health in Berry and colleagues' (2010) model is the direct effect of trauma, such as that caused by exposure to natural disasters. For example, the increase in the intensity, frequency or duration of some types of traumatic environmental events (e.g., heat waves, droughts, and heavy precipitation) that is expected to occur as part of anthropogenic climate change (see IPCC, 2013b) may result in an increased incidence of post-traumatic stress disorder. Indeed, the potential effect of climate change on the incidence of post-traumatic stress disorder has been briefly noted by the IPCC (K. R. Smith et al., 2014). Furthermore, exposure to disasters such as extreme weather events can also result in a higher risk of other types of psychopathology such as depression and anxiety (for a review see Norris et al., 2002). An increased incidence of some types of extreme weather events is one obvious and direct way that climate change may impact mental health and human behaviour. Furthermore, while Berry et al. emphasise exposure to traumatic events (such as extreme weather events) as a major direct way in which climate change may influence mental health, they also acknowledge that less acute weather conditions can also influence human behaviour and mental health. One possible such effect is an increase of aggression and violence in higher temperatures (see Anderson, 2001).

The second pathway through which climate change may affect psychological functioning in Berry and colleagues' (2010) model is via impacts on physical health. As noted previously, climate change may lead to an increase in food and water-borne diseases, heat stress, undernutrition in some populations, and other health risks (K. R. Smith et al., 2014). Berry et al. suggest that mental and physical health have a reciprocal causal relationship, and that detrimental effects of climate change on physical health may subsequently result in damage to mental health. Research using a cross-lagged panel model in a large sample of American adults indeed found evidence that poorer physical health appears to result in poorer mental health (Hays, Marshall, Wang, & Sherbourne, 1994).

The final pathway through which climate change may affect mental health that was suggested by Berry et al. (2010) is via effects of climate change on community wellbeing, for example by climate change impacting economic production. Climate change may also impact mental health and behaviour through effects on the fabric of communities by prompting migrations necessitated by rises in sea level or by other environmental problems. These migrations could in turn lead to an increased mental health burden (Page & Howard, 2010).

1.3 The Challenges of Studying the Psychological Impacts of Climate Change

The sheer willingness to engage in research concerned with the impacts of climate change is obviously a necessary requirement for psychologists to make a meaningful contribution to this research area. As mentioned previously, several recent articles (Gifford, 2008; Pidgeon & Fischhoff, 2011; Spence et al., 2009; P. C. Stern, 2011; Swim et al., 2011) have attempted to engage psychologists in climate change research. However, there has as yet been limited recognition of the fact that climate change research—and in particular, the study of its psychological impacts—presents a range of specific demands and challenges that mean that psychologists' traditional research strategies may not be well suited to this research area. In this subsection I will discuss some of these methodological challenges in more detail.

1.3.1 The challenge of timescale.

> The climatic impacts of releasing fossil fuel CO_2 to the atmosphere will last longer than Stonehenge. Longer than time capsules, longer than nuclear waste, far longer than the age of human civilization so far.
>
> Oceanographer David Archer (2009, p. 1)

In comparison to historical changes in climate, the changes to the Earth's climate currently occurring due to human activities are taking place very rapidly (see S. J. Smith, Edmonds, Hartin, Mundra, & Calvin, 2015). Nevertheless, the changes caused by current greenhouse gas emissions will unfold over a relatively long period of time (at least on a human timescale), and have consequences that may endure for centuries (Collins et al., 2013). There is therefore an intrinsic and important predictive aspect to research concerned with the impacts of climate change. Answering questions about the impacts of climate change will involve, in many cases, answering questions about the future (even the distant future). In some cases, this predictive aspect may be formally embodied in quantitative forecasts or projections. In climatology itself, the exemplars of such projections are those predicting changes in temperature over the coming decades and centuries (e.g., Collins et al., 2013). Psychologists may similarly need to produce predictions about the future effects of climate change and behaviour, with the predictions potentially applying to very long timeframes. This is a unique challenge for psychologists, who rarely engage in long-term forecasting of the behaviour of large populations. Instead, when psychologists do engage in predictions, the focus is typically on predicting the behaviour of individuals in the short term: Predictions about how an individual may change following psychological therapy, how a job applicant will perform if selected for a position, or whether a prisoner will recidivate if given parole (see W. M. Grove & Lloyd, 2006). As challenging as the making of predictions about the behaviour of individuals in the short term can be, predicting how humans may behave decades or centuries in the future may be much more difficult.

1.3.2 Novelty of the phenomenon.

This thesis began with a reminder of the climate stability throughout the geological epoch in which human civilisation emerged (the Holocene). Even in recent decades, when a clear global warming trend has become apparent, the magnitude of warming that has been observed is still much less than the warming likely over the remainder of this century. Mean global surface temperatures increased by about 0.85°C between 1880 and 2012 (IPCC, 2013b), but the warming expected over the 21st century is much more substantial. As mentioned previously in section 1.1.2, if future greenhouse gas emissions follow the trajectory assumed in the RCP8.5 scenario (rising emissions throughout the 21st century), warming of between 2.6 to 4.8°C is likely to eventuate. This acceleration in warming is due to the rapidly increasing concentration of greenhouse gases in the atmosphere. Recently, a carbon dioxide concentration of over 400 parts per million was detected for the first time at the Mauna Loa observatory, a concentration that is likely higher than the Earth has experienced in the last three million years (Gillis, 2013).

To help make predictions about climate change despite the novelty of events such as the rapid recent increase in greenhouse gas concentrations, paleoclimatologists and geologists have access to indirect measurements of important climatic variables such as global temperatures and greenhouse gas concentrations stretching back several hundred million years. These paleoclimatological temperature records are based on information gathered from ice cores, tree rings, sediments, and other proxy sources of climatological measurements (see National Climatic Data Center, n.d.). These measurements can be used to identify past events where changes in greenhouse gas concentrations and/or temperatures were similar to those currently occurring. Studying these past events can in turn help physical scientists to predict the ecological consequences of contemporary climatic change. The Paleocene-Eocene Thermal Maximum, a period around 55 million years ago during which the Earth’s temperature increased very rapidly, is one useful analogue for modern climate change (Webb, Leighton, Schellenberg, Landau, & Thomas, 2009).

Researchers interested in climatic impacts on human behaviour do not have the luxury of records showing how global climatic changes similar to those likely to occur over the next century have impacted humans in the past. As mentioned previously, modern humans have only existed for around 200,000 years, and human civilisation has existed for only a few thousand years. Modern human society has simply never faced the types of global climatic changes that will occur over the next century. This means that psychologists do not have access to detailed observational data demonstrating the relationships between human behaviour and global climate change. It is possible to correlate short-term global and regional meteorological variations with social and behavioural variables, but extrapolating such findings to forecast the impact of future climate change requires the very significant assumption that long-term climatic changes will have similar effects on humans as do short-term weather variations. This assumption simply may not hold in practice. The novelty of the problem of climate change—and the resulting lack of data showing how humans have responded to major changes in climate in the past—is a major challenge for psychologists hoping to predict how human behaviour will respond to climate change.

1.3.3 Variability of impacts.

Different human populations may also be affected by climate change in very different ways. For example, the populations of small island nations may have their habitats threatened by sea level rise (Lewis, 1990), while populations living at higher altitudes will not. Furthermore, even the effects of exposure to the same change in local environmental conditions may vary across individuals and populations. For example, a study of heat-related mortality in the subtropical Australian city of Sydney (Vaneckova, Hart, Beggs, & de Dear, 2008) found a smaller relationship between heat and mortality than in temperate developed countries. Vaneckova et al. suggested that those living in subtropical climates may display behavioural, technological, and architectural adaptations to warmer weather, meaning that a heat wave involving the same temperature might be less dangerous in a subtropical region than in a temperate one. Some populations are certainly more vulnerable to climate change than others: Specifically, populations that are socially, politically and economically marginalised are more vulnerable to climate change than those who have greater physical and social resources (IPCC, 2014b). Because the impacts of climate change on behaviour and psychological variables may differ across populations, psychologists interested in the effects of climate change may need to study diverse human populations in order to avoid producing conclusions that hold only for a restricted subset of humans. Some more subtle implications of the possibility that the effects of climate change may vary across time and place are discussed later in this thesis (see section 5.2.3.2.1).

1.3.4 Communication with policymakers, physical scientists and non-scientists.

Finally, psychological research concerned with climate change will often need to be communicated effectively to policymakers and other members of society who are not themselves psychologists. Such communications may be necessary for a variety of reasons. One reason is that the scale of climate change as a practical problem means that attempts to mitigate or assist humans in adapting to its effects may require governmental or inter-governmental support to be applied effectively, meaning that the communication of findings to policymakers may be required.

A second reason why psychological research concerned with climate change will often need to be communicated effectively to non-psychologists is that the study of climate change is an interdisciplinary field. The integration of findings from the physical sciences, life sciences, and social sciences is crucial in order to provide a clear picture of the causes and consequences of climate change. The most well-known venue for the integration of climate change research is in the reports of the IPCC, which are prepared (for the most part) by physical scientists. Psychological research that is presented in such a manner that it does not reach, cannot be understood by, or is dismissed by physical scientists may not appear in integrating work such as IPCC reports, and have diminished impact on policymaking and other practical responses to climate change.

A final reason why effective and careful communication of findings to non-psychologists may be necessary is that climate change is a politically charged topic that is paid a great deal of attention by some segments of the general public (including environmentalists, climate change "sceptics", and other interested groups). The findings of climate change research may be disseminated by pro-environmental groups and used as the basis for political action. Conversely, research reports on the topic of climate change may face intense and even threatening scrutiny from those who are sceptical of the hypothesis that human activities are responsible for warming of the global climate (see M. E. Mann, 2012).

The task of reporting findings to non-psychologists is of course a common problem for psychologists engaging in applied research. The special challenge that climate change presents is that of communicating with groups that are not just unfamiliar with psychologists' methods and theories, but that may also expect a very high standard of research (e.g., interdisciplinary climate scientists), be aggressive toward and dismissive of research in this area (e.g., climate change "sceptics"), or that may only make practical use of research if its practical implications are communicated very directly and persuasively (e.g., governments).

1.4 The Methodological Readiness of Psychologists to Study the Impacts of Climate Change

Studying the potential effects of anthropogenic climate change on human behaviour and psychological functioning is a unique problem for psychologists. It is a problem that requires long-term predictions about the effects of a phenomenon that is novel within recorded human history and that represents a grave and complex change to humans' physical environments. It is therefore important to consider whether psychologists are prepared to deal with the methodological challenges that research in this area presents. How suitable are the mainstream methodological strategies of psychological researchers for studying the potential impacts of climate change? And how can psychologists' methodological strategies best be applied or altered to meet the challenges of this potentially treacherous research area?

The broad objective of this thesis, then, is to investigate how well the mainstream methodological strategies used in psychology are suited to the challenges involved in studying the psychological impacts of climate change. The term "methodological strategies" implies a focus on broad conceptual issues in methodology, such as the range and nature of the sources of uncertainty accounted for in psychological studies, and the broad types of statistical inference used. I will focus less on specific research tactics, such as the methods used to estimate statistical models or to conduct specific experiments. As alluded to above, I will focus also on more commonly applied mainstream methodological strategies in psychology, restricting the focus to methods that are broadly quantitative in nature. This restriction in focus is not intended to imply that less commonly used or qualitative methods are not of interest. Rather, the intent is to produce a critique that is of general a scope and broad in applicability as possible while not being overly cumbersome in length.

Such a methodological critique obviously needs to draw on existing psychological literature related to the potential effects of anthropogenic climate change on human behaviour. Importantly, climate change may exert effects on human behaviour via effects on mediating variables (such as temperature or the frequency of extreme weather events) whose effects may be studied in many psychological research projects, not just those explicitly concerned with climate change. The findings of such projects could certainly be of some relevance when drawing inferences about climate change. This implies that a simple examination of psychological literature explicitly mentioning climate change might well exclude a number of studies that could be informative as to psychology's methodological readiness to produce effective and useful research concerned with the impacts of climate change. On the other hand, the breadth and diversity of human-climate interactions means that the overarching body of psychological research that is potentially relevant when drawing inferences about climate change impacts would be so large and diverse that the performing of a coherent methodological critique of this entire body of literature would be a nearly intractable challenge.

1.5 The Testbed: Temperature-Violence Research

Therefore, rather than attempting to review all psychological research that is explicitly concerned with the impacts of climate change, or that is potentially relevant when making inferences about such effects, I will undertake a methodological critique drawing on one specific area of psychological research that is relevant to the study of the impacts of anthropogenic climate change. This area is that of research concerned with the effects of temperature on the incidence of intra- and interpersonal violence. There is an extensive current body of literature dealing with the effects of temperature on both the incidence of acts of interpersonal violence such as assault and homicide (see Anderson, 2001 for a review) as well as on the incidence of acts of *intra*personal violence such as self-harm and suicide (for reviews see Deisenhammer, 2003; and Dixon & Kalkstein, 2009)[2]. This area of study has several key features that make it

[2] There is also a fairly broad literature concerned with the relationship between temperature and acts of inter-*group* violence (e.g., wars and civil conflicts; see Hsiang, Burke, & Miguel, 2013). I do not discuss this research area in much detail in this thesis, given that it falls outside of the traditional domain of psychology.

suitable for examining psychology's methodological readiness for climate change research.

Firstly, the study of the relationship between weather and violence has a long history in psychology and related fields such as sociology (see Durkheim, 1897; Leffingwell, 1892; Morselli, 1882), and continues to be a popular object for both theoretical and empirical psychological study (e.g., Anderson & Anderson, 1998; Anderson & DeLisi, 2011; Bushman, Wang, & Anderson, 2005; Page, Hajat, & Kovats, 2007; Preti, Lentini, & Maugeri, 2007; Rotton & Cohn, 2003). Much of this research is not explicitly concerned with climate change, but is nevertheless of potential relevance to questions about climate change. The breadth and long history of literature concerned with temperature effects on violence means that there is a wide range of studies in this area that can be drawn on when evaluating the suitability of psychologists' methodological strategies for studying the impacts of climate change.

Secondly, despite the presence of many studies studying the relationship between temperature and violence that have not explicitly recognised the connection of this topic to climate change, there have been a number of studies that *have* highlighted this connection (e.g., Anderson, 2001; Anderson, Anderson, Dorr, DeNeve, & Flanagan, 2000; Anderson & DeLisi, 2011; Gamble & Hess, 2012; Helama, Holopainen, & Partonen, 2013; Preti et al., 2007). The fact that at least some researchers in this area are directly grappling with the issue of climate change makes this research area a particularly appealing choice as a testbed in which to identify and illustrate methodological problems and issues. Some researchers in this area have even attempted to predict the quantitative effect of future climate changes on behavioural variables (e.g., Anderson, 2001; Anderson et al., 2000; Anderson & DeLisi, 2011; Rotton & Cohn, 2003). This is an especially important and unusual feature of this research area, and one that will be helpful when discussing the methodological issues related to such predictions.

Thirdly, research concerned with the effects of temperature on intra- and interpersonal violence is relatively representative of mainstream psychological research practice in several respects. It is a research area that is primarily quantitative in orientation; that employs both laboratory experiments (e.g., Anderson et al., 2000; Vrij, Steen, & Koppelaar, 1994) and correlational designs (e.g., Anderson & DeLisi, 2011; Lester, 1999); that uses popular data analytic strategies in psychology such as analysis of variance (ANOVA; e.g., Anderson et al., 2000), regression (e.g., Rotton & Cohn, 2003), and null hypothesis significance testing; and whose findings are often reported in mainstream psychological journals such as the *Journal of Personality and Social Psychology* (e.g., Anderson & Anderson, 1996; Anderson, Bushman, & Groom, 1997), and *Psychological Science* (e.g., Anderson, Benjamin, & Bartholow, 1998). These characteristics make this research area a suitable testbed with which to tentatively draw conclusions about the strengths and weaknesses of mainstream methodological practice in psychology when applied to the study of climate change impacts.

The methodological critique contained within this thesis will therefore draw on published psychological literature concerned with the relationship between temperature and both intra- and interpersonal violence as a key source of information about psychology's methodological readiness to engage in climate change research. This section of literature does include research reports concerned with topics of traditional interest to psychologists but written by authors not necessarily identifying as psychologists (e.g., psychiatrists and criminologists); for the purposes of the review the term "psychological research" will be interpreted in a reasonably flexible manner. Furthermore, while the testbed of research concerned with the relationship between temperature and violence will be given particular focus, examples of psychological studies that do not fall within this testbed—especially those focusing on the impacts of climate change—will also be discussed where they are particularly relevant.

1.5.1 Applied engagement with temperature-violence research.

A review and critique of the literature concerned with the relationship between temperature and violence will not be the only way in which this research area will inform my methodological critique. Actually engaging in research frequently raises issues not obvious to the observer passively reading the published outputs of a particular research area. It also allows for the examination of methodological issues that have rarely presented themselves within the published literature, but which may do so in the future. For example, engaging in applied research permits a more detailed examination of the problem of producing quantitative predictions than would be possible given the existing but limited number of instances of psychologists producing quantitative predictions about climate change impacts. As well as drawing on published psychological literature concerned with temperature and violence, I will therefore also engage in an applied example of this research.

The applied component will take the form of three empirical studies investigating the relationship between temperature and the incidence of assaults, suicides, and self-harm hospitalisations in New Zealand. This research was conducted with the specific goal of determining what the past relationship between temperature and violence in New Zealand can tell us about the likely effects of future climate change on intra- and interpersonal violence. However, it is crucial to stress that while the empirical component of a research study would traditionally be the centrepiece of a PhD project, this is not the case for my thesis. The applied component is rather positioned as an information source for the central object of my thesis (the methodological critique in section 5). The applied component will be presented in the form of three journal articles. The articles deal with the effects of temperature on assault, suicide, and self-harm respectively.

1.5.2 Contemporary methodological literature.

As well as the testbed of research concerned with temperature and violence, contemporary methodological literature will be drawn on as necessary when evaluating the suitability of particular methodological strategies to the study of climate change impacts. This thesis will represent the first major attempt at a discussion of the methodological problems arising in psychological climate change research. However, existing methodological literature concerned with the behavioural sciences more generally will be an important source of information when judging the suitability of particular research strategies to the study of climate change impacts. Of particular interest will be literature concerned with broad methodological issues in psychology. These issues include whether our framework of statistical inference should be frequentist or Bayesian (e.g., Wagenmakers, 2007); whether psychological attributes are actually quantitative (Michell, 1997, 2000); and how psychologists should report the size of effects (e.g., Baguley, 2010; Cumming et al., 2007). Literature concerned with these broad methodological issues may be of particular value in the context of psychological climate change research: The novel and severe demands of this research area may force psychologists to make extensive and structural changes to how we go about doing research. Furthermore, the consideration of broad methodological issues common to many psychological research projects may be of assistance in producing a methodological critique that, while focused on a specific testbed of research, is nevertheless relevant to psychological research concerned with the impacts of climate change more broadly.

1.5.3 Limitations of the testbed approach.

In limiting my critique of psychology's methodological readiness to deal with climate change research to a focus on a specific testbed of psychological research, the breadth of research to be considered is necessarily limited. One specific (and intentional) limitation is that this project will be focused on research concerned with the potential psychological *impacts* of climate change, as opposed to research concerned with *mitigating* climate change. This decision does not seem unjustifiable: Although psychological research concerned with climate change mitigation is valuable, research concerned with the psychological effects of climate change is important for the reasons discussed earlier in section 1.2.2: At least some global warming over the following century is inevitable, and the effects on humans could be severe. While the methodological challenges arising when engaging in psychological research concerned with the mitigation of climate change are certainly also worthy of discussion, for reasons of practicality this thesis is restricted in focus to the study of the impacts of climate change.

More broadly, no individual research area is completely representative of the entire discipline of psychology, either with respect to its methodological practices or any other features. This means that my capacity to draw conclusions about psychology's readiness as a discipline to deal with the methodological challenges involved in studying the impacts of climate change will necessarily be somewhat limited. However, by critically discussing the specific methodological issues that arise within this section of research, I aim to be able to provide criticisms and suggestions that are relevant to a broad range of psychological research projects concerned with the effects of climate change.

1.6 Conclusion

The Earth's climate is changing, and it is changing due to human activities. Human behaviour causes climate change, but human behaviour will also be *affected* by climate change. Psychologists are therefore beginning to study the potential impacts of climate change on human behaviour and psychological functioning. This research area presents several significant methodological challenges. These challenges include the long timescale of climate change, the diversity of its effects, and its sheer novelty as a problem for modern humans. The major aim of this thesis is to evaluate psychology's readiness to deal with the methodological challenges involved in studying the impacts of climate change. In pursuing this aim I will draw on three major sources: an area of psychological literature relevant to anthropogenic climate change (literature concerned with the effects of temperature on intra- and interpersonal violence); an applied engagement in psychological climate change research; and a range of contemporary methodological literature. In the following three chapters, I present three empirical studies comprising an applied engagement in research concerned with the potential effects of climate change on human behaviour. These articles are focused on the effects of temperature on assault, suicide, and self-harm respectively. Each article includes a review of the literature relevant to the study, meaning that I have not included a separate review of the literature concerned with temperature and violence. Following these three empirical articles, the core component of this thesis—the methodological critique—will be presented.

2 Study One: The Relationship between Temperature and Assault in New Zealand

The following article has been published by *Climatic Change.* The version shown is the accepted manuscript, with minor formatting changes for consistency with the rest of this thesis, and some typographical errors corrected. An electronic supplementary materials document that was provided along with the main text is provided in Appendix A (section 9), and contains additional technical details about the study's method and results. References for this article are included within the consolidated reference list in section 8. The final publication is available at Springer via the citation below.

Williams, M. N., Hill, S. R., & Spicer, J. (2015). The relationship between temperature and assault in New Zealand. *Climatic Change*, *132*(4), 559–573. http://doi.org/10.1007/s10584-015-1438-7

2.1 Abstract

A number of previous studies have reported a positive relationship between ambient temperature and the incidence of violent crimes such as assault. This has led some authors to suggest that anthropogenic climate change may result in an increase in violent crime rates. In this study, we report an investigation of the relationship between temperature and assault incidence in New Zealand. Both police data listing recorded assaults as well as data from the Ministry of Health listing hospitalisations due to assault were examined. Geographical, seasonal, and irregular daily variation in temperature were all positively related to the incidence of assault, although only the effect of irregular variation in temperature was robust to controls for plausible confounds. The estimated effect of irregular daily variation in temperature was approximately 1.5% extra recorded assaults for each 1°C increase in temperature. It remains difficult, however, to make accurate predictions about future assault rates in a warming world. For example, humans may react to sustained changes in climate in ways that differ markedly from their reaction to short-term variation in temperature. Climate change may also affect rates of violence via mechanisms other than those that currently drive the relationship between temperature and violence. Furthermore, assault rates may continue to change in response to factors unrelated to climate change, such as those responsible for the long-term historical decline in human violence.

2.2 Introduction

A number of previous studies have found a positive relationship between temperature and the incidence of acts of interpersonal violence such as assault (see reviews by Anderson, 2001; Anderson et al., 2000). If increasing temperatures precipitate violence, then anthropogenic climate change might be expected to result in higher rates of violence (Anderson & DeLisi, 2011). This research area is therefore one of potential interest to researchers concerned with the impacts of climate change.

2.2.1 Geographical variation in temperature and interpersonal violence.

One way to study the relationship between temperature and the incidence of interpersonal violence is to compare hotter and colder geographical regions. A positive relationship between geographical variation in temperature and homicide rate was found across countries in two international studies (Robbins, Dewalt, & Pelto, 1972; Rotton, 1986). Anderson et al. (2000) also found that temperature was a significant predictor of violent crime (with murder, rape and assaults as indicators) across 260 US cities. This relationship had a standardised effect size of 0.43 when controlling for socio-economic status, population level, and sociocultural “Southernness”.

2.2.2 Temporal variation in temperature and interpersonal violence.

Another way to study the relationship between temperature and interpersonal violence is to examine the relationship between the incidence of interpersonal violence and variation in temperature *within* a region (temporal variation in temperature). Temporal variation in temperature can be broken down further into sub-components including seasonal and irregular (random) variation in temperature. In terms of seasonal variation in temperature, Anderson et al. (2000) reported a comparison of assault rates across seven Northern Hemisphere datasets, finding that monthly assault rates were higher in warmer months: more than 40% higher in the hot month of August than in the cold month of January. Hipp et al. (2003) also found that violent crime rates in the United States were about 35% higher at the summertime peak than at the winter nadir.

Another type of temporal variation in temperature is *irregular* variation: random variation, not explained by regular seasonal fluctuations. Irregular variation in temperature can be studied in several different ways. One method is to aggregate data (e.g., at the annual level), such that the effects of seasonal variation are removed. A series of studies using this approach and data from the United States found a positive relationship between temperature and violent crime (Anderson et al., 2000, 1997; Anderson & DeLisi, 2011; Rotton & Cohn, 2003). Another way in which irregular variation in temperature can be examined is by using data at a finer temporal level of aggregation (e.g., by using days or even shorter intervals), but using temporal control variables (e.g., for the months of the year) to remove seasonal variation. Positive effects of temperature on assault incidence have also been found by studies using this approach (Cohn & Rotton, 1997; Harries & Stadler, 1983).

A meta-analysis of the effect of temperature on interpersonal violence was conducted by Hsiang, Burke, and Miguel (2013). Hsiang et al. ensured that each study included in their meta-analysis controlled for any differences between geographical areas (if more than one region was included within a study), and for the effect of time, meaning that their study focused on the effect of irregular variation in temperature. In some cases, these controls were applied in re-analyses of the original studies' data. Their meta-analysis found a mean effect of a 2.3% increase in interpersonal violence for each standard deviation increase in temperature.

2.2.2.1 Experimentally manipulated temperature

As well as the aforementioned field studies, some laboratory studies have also found an effect of temperature on aggressive behaviour. For example, a Dutch study found that police officers were more likely to draw weapons and shoot suspects in a simulated burglary when the temperature in the simulation room was hotter (Vrij et al., 1994). However, the effect of temperature on aggressive behaviour in experiments is not consistent enough to result in a significant effect when synthesised in a meta-analysis (Anderson et al., 2000). One explanation for the inconsistent findings in this area is that the mechanism via which temperature affects violence in the real world cannot operate in the laboratory: for example, if temperature increases precipitate violence by providing increased opportunities for interactions between potential offenders and victims (Cohn & Rotton, 2000).

Given the lack of firm experimental evidence, it is worth considering whether the generally positive relationship between temperature and violence could be attributable to some confounding variable. The fact that an apparent effect of temperature has been observed regardless of whether the component of temperature variation examined is geographical, seasonal or irregular greatly reduces the list of plausible confounds. Cultural or demographic variables might explain a relationship between geographical variation in temperature and violence (as in Hipp et al., 2003), and changes in human activity patterns due to calendar events (e.g., increased vacationing in summer) might explain the relationship between seasonal variation in temperature and violence. However, neither of these potential confounds can explain why irregular variation in temperature is related to violent crime rates. Overall, a causal effect of temperature on violence is probably the most plausible reason for the relationship between temperature and interpersonal violence, although the specific mechanism of this relationship is subject to much uncertainty.

2.2.3 The mechanism of the temperature-violence relationship.

A number of theories have been invoked to explain a relationship between temperature and the incidence of interpersonal violence. The general affective aggression model (Anderson, Deuser, & DeNeve, 1995) suggests that uncomfortable temperatures can increase aggressive cognitions and emotions and thereby increase the incidence of aggressive behaviour. Similarly, the negative affect escape model (Bell & Baron, 1976; see also Bell & Fusco, 1989) suggests that unpleasant situational variables such as uncomfortable temperatures can increase negative affect, and that negative affect increases aggressive behaviour. However, in the negative affect escape model this relationship is hypothesised to hold only up to an (unspecified) inflection point, beyond which extremely high levels of negative affect prompt escape from the situation. Whether non-linearity in the temperature-assault relationship exists such that very high temperatures prompt a decrease in assaults is a subject of some debate (Bell & Fusco, 1989; Bushman et al., 2005; Cohn & Rotton, 2005).

An alternative sociological theory is that warmer weather precipitates violent crime by increasing the probability that potential offenders and victims come into contact (Cohn & Rotton, 2000). Finally, it is possible that warmer temperatures provoke violence by eliciting greater alcohol consumption (Harries & Stadler, 1988). Indeed, alcohol sales tend to be highest in summer (Uitenbroek, 1996), and there is strong evidence that alcohol consumption increases the risk of aggressive and violent behaviour (Bennetts & Seabrook, 2008; Bushman & Cooper, 1990; Ray et al., 2008).

2.2.4 Consideration of climate change.

Most studies concerned with the relationship between temperature and interpersonal violence have been completed without an explicit consideration of what their findings imply for a future that involves a warming climate. One notable study that did pay explicit attention to climate change was Anderson and DeLisi (2011; see also similar earlier studies by Anderson et al., 1997, 2000). Anderson and DeLisi correlated national annual rates of serious and deadly assault in the United States for 1950–2008 with temperatures averaged across the 50 largest cities in the country. The authors estimated that there were an extra 7.54 murders and assaults per 100,000 for each 1°C temperature increase when controlling for year, autoregressive effects, and the proportion of the population in prison. They reported that this implied that an 8°F (4.4°C) temperature increase would lead to more than 100,000 extra serious and deadly assaults per year in a population of 305 million.

Another study that did explicitly consider climate change was Cohn and Rotton (2003). Cohn and Rotton analysed both US national data as well as cross-sectional time series assault data aggregated at state level (for 1960–1998). On the basis of this analysis they calculated a smaller effect of an extra 2.18 additional assaults per 100,000 p.a. for each 1°C temperature increase. Finally, using a series of temperature anomalies from St. Louis, Missouri, Mares (2013) found that each 1°F increase in temperature was associated with a 0.69% increase in violent crime (1.25% per 1°C), and concluded that climate change may have a substantial effect on crime rates.

Of course, anthropogenic climate change is likely to comprise not just temperature increases but changes to other meteorological variables such as precipitation, wind, and tropical storm intensity. In this study we focus on the effects of temperature, primarily because future changes to other meteorological variables are subject to more uncertainty and are expected to differ widely across geographical regions (see Reisinger et al., 2014 for projections for New Zealand). The large body of existing literature concerned with the relationship between temperature and violence also provides a basis for a concern that temperature increases in specific might have a detrimental effect on assault incidence. This said, we also briefly examine the effects of humidity, given the effect of humidity levels on thermoregulation and apparent temperature for humans.

2.2.5 Research aims.

In this study, we aimed to investigate the relationship between geographical, seasonal, and irregular variation in temperature and the incidence of one important type of interpersonal violence—assault—in New Zealand. In order to further assess the plausibility of the hypothesis that global meteorological changes can affect assaults in New Zealand, we also examined the relationship between the El Niño Southern Oscillation (ENSO) and assaults in New Zealand. We aimed to use the information produced in our analyses to predict how rising temperatures may affect the incidence of violence in New Zealand in the coming decades. We present investigations of two separate datasets relating to the incidence of assault in New Zealand: recorded assault data from the NZ Police, and assault hospitalisation data from the Ministry of Health.

Like those of most other studies in this research area, our empirical analyses are restricted for practical reasons primarily to observations about the relationship between human behaviour and the *weather* (i.e., relatively short term variation in meteorological conditions, particularly temperature). Our overarching goal is nevertheless to make tentative inferences about the effects of *climate* change on assaults (the term climate referring to the mean and variability of meteorological variables over a long period of time, classically 30 years; Planton, 2013). Such inferences are necessarily subject to significant uncertainty: Humans may react differently to sustained climate changes than they do to short-term temperature variation.

2.3 Methods

2.3.1 Recorded assault data.

Data listing daily recorded assaults from 1 July 1994 to 31 July 2009 within five police districts (Auckland City, Waitematā, Counties Manukau, Wellington, and Canterbury) were obtained from the New Zealand Police. The Auckland City, Waitematā, and Counties Manukau police districts all cover different areas of the wider Auckland region, and were thus combined into a single Auckland region.

2.3.2 Assault hospitalisation data.

The New Zealand Ministry of Health provided a listing of all public hospital admissions in New Zealand with a discharge date between 1 January 1993 and 31

December 2009 where at least one of the causes of hospitalisation was assault. These data were collated by date of injury (not necessarily the date of admission) and the patient's territorial authority area of domicile. Territorial authority areas are hereafter referred to as "districts". Excluding the offshore Chatham Islands, there are currently 66 districts in New Zealand. Where assaults on multiple dates contributed to a hospitalisation, the most recent date of assault was utilised as the date of injury. Furthermore, where the same assault appeared to have caused multiple hospitalisations (i.e., when the patient and the most recent date of injury listed for two hospitalisations matched), only the first admission was counted.

The original source of the hospitalisation data was the national minimum dataset for hospital events. A problem with this dataset was the presence of inconsistencies in terms of whether or not short emergency department stays were recorded, with different reporting practices used by different district health boards, as well as changes in reporting practices occurring over time (see Ministry of Health, 2012 for a discussion). We therefore followed the practice of the Ministry of Health in excluding emergency department stays of less than two days from analysis, resulting in the exclusion of 19,356 incidents. A further 703 hospitalisations with a district of "overseas/other" listed were excluded, as were five hospitalisations of patients living on the offshore Chatham Islands.

2.3.3 Population estimates and demographic data.

Annual population estimates by police area and by district were obtained from Statistics New Zealand and interpolated to produce daily estimates. Ethnicity and age data by district were obtained from Statistics New Zealand as at the census years 1996, 2001, and 2006. Deprivation index (NZDep) values by area for the same three censuses were obtained from the University of Otago (Salmond, Crampton, Sutton, & Atkinson, 2014). NZDep values were aggregated into a single average value for each district over these three time points. The NZDep is a measure of socioeconomic deprivation scaled to have M = 1000 and SD = 100.

2.3.4 Meteorological data.

Meteorological data was obtained from the virtual climate network of the National Institute of Water and Atmospheric Research (NIWA, n.d.-b). This network uses a thin plate smoothing spline model to interpolate meteorological measurements on a regular 5km grid across New Zealand (see Tait, Henderson, Turner, & Zheng, 2006). Use of the virtual climate network avoided any problems with missing data or changing station locations that could arise from the use of physical stations. The virtual station used to represent each region or district was the closest station to the town centre of the largest town within that region or district. Southern Oscillation Index (SOI) values were obtained from the Australian Bureau of Meteorology (n.d.). The SOI is a measure of ENSO variability based on the seasonally adjusted mean sea level pressure difference between Tahiti and Darwin.

2.3.5 Ethical approval.

The study received ethical approval from the Massey University Human Ethics Committee.

2.3.6 Data analysis and computation.

Data analysis was completed using R version 3.1.1 (R Core Team, 2013). Generalised linear mixed models were estimated using the function glmmPQL in the MASS package version 7.3-34 (Ripley, 2014). A Poisson distribution and log link were used for almost all the models reported. glmmPQL includes a freely estimated residual variance term when estimating Poisson models, avoiding any problems with overdispersion.

2.3.7 Electronic supplementary materials.

Further information about this study's methods and results is available in the electronic supplementary materials for this article.

2.4 Results

2.4.1 Analyses of assaults recorded by the police.

In this first subsection of the results, analyses of assaults recorded by the police are reported. Over the study period, there were 162,219 recorded assaults in Auckland, 46,838 in Canterbury, and 58,958 in Wellington (combined total 268,015). The mean number of recorded assaults per 100,000 was 846 p.a. in Auckland, 891 in Wellington, and 616 in Canterbury (combined mean 803 per 100,000 p.a.). The mean temperature at the city centre of each region was 15.3°C in Auckland, 13.2°C in Wellington, and 12.0°C in Canterbury.

2.4.1.1 Form of the temperature-assault relationship.

Scatter plots displaying the relationship between temperature and assault rates are displayed in Figure 2. Loess smoothing was used to indicate a line of best fit in each case. These plots indicate firstly that the relationship between temperature and assault (for both recorded assaults and assaults resulting in hospitalisation) is relatively small in comparison to the wide variation in daily assaults. Secondly, there is little evidence of any marked deviation from linearity: The small positive relationship between temperature and assault held regardless of whether temperatures were cold, moderate, or hot. Linear or log-linear relationships were therefore assumed in the models estimated in this study. The estimated effects of seasonal and irregular daily variation in temperature within each region are reported in the subsections that follow.

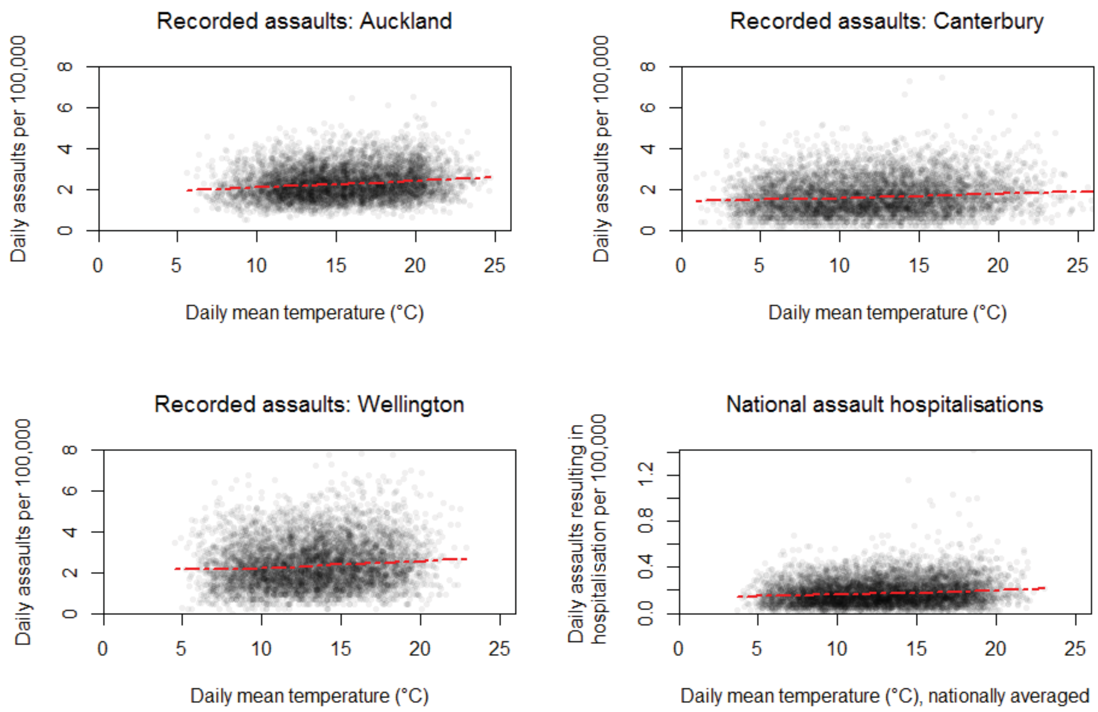

Figure 2. Scatter plots of temperature and recorded assaults in Auckland, Wellington, and Canterbury, and of temperature and assaults resulting in hospitalisation (nationally averaged). Lines of best fit shown using loess smoothing (span = 0.5, degree = 2).

2.4.1.2 Seasonal variation in temperature and recorded assaults.

In all three regions, assaults were least frequent in June (winter), as is visible in Figure 3. The peak month was December for both Auckland and Wellington, and November in Canterbury.

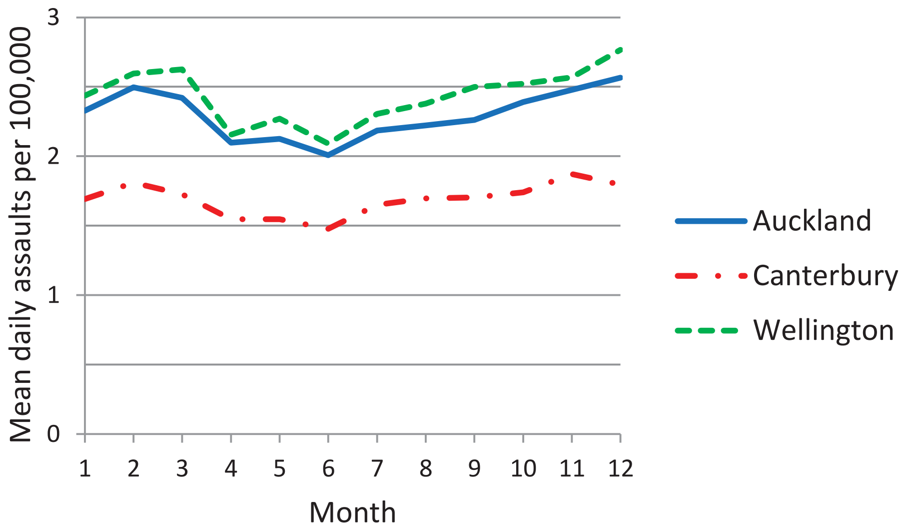

Figure 3. Mean daily recorded assaults by month and region.

The effect of seasonal variation in temperature on recorded assaults was estimated using a generalised linear mixed model estimated within each of the three regions using a Poisson distribution. In this model, the predictor variable of interest was the mean temperature on each of the 365 days of the calendar year as averaged across the entire study period within each region (i.e., the seasonal norm temperature). The response variable was the summed number of assaults occurring in a region across each of the 15 occurrences of each day of the calendar year over the study period. In order to ensure that each day of the year had the same number of instances, the period analysed was restricted to 1 August 1994–31 July 2009, and February 29 was excluded. The intercept was free to vary randomly across regions, removing any effect of geographical variation in temperature. The effect of population was included as a control variable, following a log transformation; in combination with the log link function, this allowed population size to have an additive effect. One autoregressive term was specified for the error terms in order to deal with short-term residual autocorrelation.

When entered as the sole predictor apart from population, the coefficient for seasonal variation in temperature was positive, $\hat{\beta}$ = 0.014, 95% CI [0.011, 0.017].

However, if daylength was controlled as an additional predictor, the estimated effect of temperature fell to nearly zero, $\hat{\beta}$ = -0.004, 95% CI [-0.008, 0.0003]. The log of daylength (in hours) had a larger effect of $\hat{\beta}$ = 0.461, 95% CI [0.378, 0.543], implying that every 10% increase in daylength was associated with a $1.10^{0.451} - 1$ = 4.5% increase in assaults. Seasonal variation in temperature and daylength were correlated an average r = .79 across regions.

2.4.1.3 Irregular daily variation in temperature and recorded assaults.

Irregular daily variation in temperature was captured by calculating the difference between the temperature observed in a region on a given date and the seasonal norm temperature (as defined above). Across dates and districts, the standard deviation of these temperature anomalies was 2.2°C, with a range of -8.0 to 8.7°C. The effect of irregular temperature variation was estimated again using a generalised linear mixed model, with a Poisson distribution, and population size controlled. Intercepts were free to vary across both regions and years (within regions), allowing for a flexible control for any omitted time-varying confounders. The response variable was simply the number of assaults occurring on a given date within a region. Dummy variables for month and weekday and a single autoregressive term for the error structure were included in order to account for residual autocorrelation. In this model, the estimated effect of temperature was $\hat{\beta}$ = 0.015, 95% CI [0.012, 0.017], or around 1.5% more assaults for every 1°C increase in temperature. This coefficient also implies that a standard deviation (2.18°C) increase in temperature anomaly was associated with an $e^{0.015*2.18} - 1$ = 3.3% increase in assaults.

If the effect of temperature was permitted to vary across regions, it was very similar in each case: $\hat{\beta}$ = 0.014 in Auckland and Canterbury, and 0.017 in Wellington. The lagged effects of temperature anomalies on recorded assaults were also investigated with lags up to 7 days in an alternative model, but all of the lagged effects were very small ($|\hat{\beta}| < 0.003$) and not statistically significant. When the relative humidity anomaly (%) was added to the model along with temperature, it had a very small negative main effect $\hat{\beta}$ = -0.001, 95% CI [-0.001, -2.3×10^{-4}], while there was no evidence of an interaction between relative humidity and temperature, $\hat{\beta} = 2.0 \times 10^{-4}$, 95% CI [$-1.4 \times 10^{-5}$, 4.1×10^{-4}].

2.4.1.4 Trends in recorded assaults and temperature.

Year-to-year trends in recorded assaults and mean temperature (averaged across the three regions studied) are displayed for descriptive purposes in Figure 4. There was a slight positive trend in assault rate, with an increase of 3.2 extra assaults per 100,000 per year.

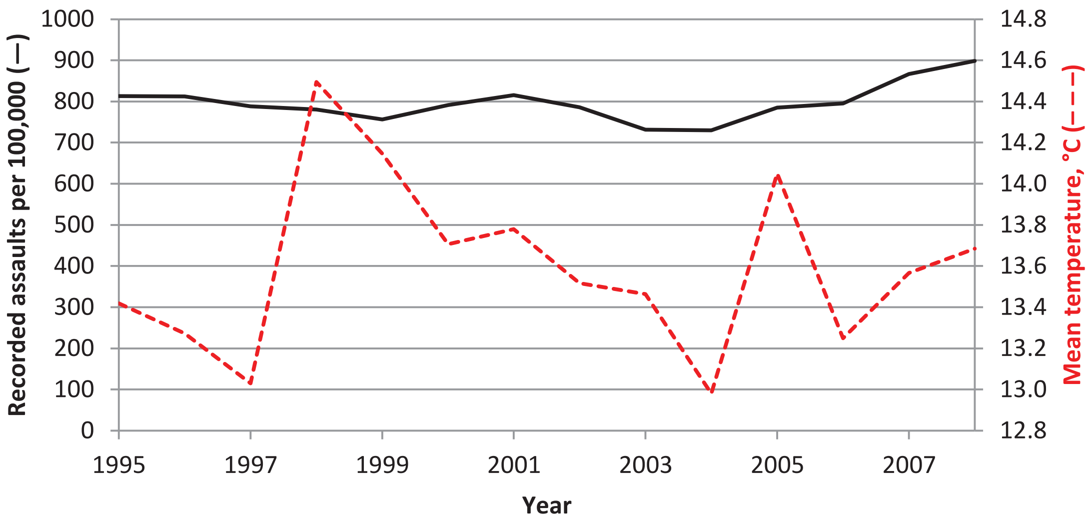

Figure 4. Trends in recorded assaults and temperature; averaged across Auckland, Wellington, and Canterbury.

2.4.1.5 Recorded assaults and ENSO.

High SOI values indicate La Niña conditions, which tend to be associated with warmer temperatures in most of New Zealand (Wratt, Basher, Mullan, & Renwick, n.d.). During the study period, the correlation between SOI values and monthly mean temperatures (averaged across Auckland, Wellington and Canterbury) was r = .17. The correlation between monthly SOI values and the mean number of recorded assaults per 100,000 per day (aggregated across Auckland, Wellington and Canterbury) was r = .24, 95% CI [.10, .37]. A linear mixed model was also estimated with the sum of assaults across Auckland, Wellington and Canterbury in each month as the response variable, a random effect for year to control for any omitted time-trending confounds, and a Poisson distribution with log link. In this model the coefficient for SOI remained positive, $\hat{\beta}$ = 0.002, 95% CI [0.001, 0.004]. Given that SOI values are scaled to have a long-term standard deviation of 10, this coefficient suggests a 2% increase in assaults for every 1σ increase in SOI.

2.4.2 Analyses of hospitalisation data.

In this second subsection of the results, we report analyses based on nationwide assault hospitalisation data. The use of nationwide data allowed for the study of the effects of geographical variation in temperature as well as of irregular and seasonal variation. There were a total of 44,745 assaults resulting in hospitalisation over the study period of 1993 to 2009, or 67 incidents per 100,000 p.a. As mentioned previously, this dataset excludes short emergency department stays. It is nevertheless clear that the rate of assault resulting in hospitalisation is much smaller than that of recorded assaults. There was a moderately strong correlation, r = .52, between the number of daily recorded assaults across the Auckland, Wellington and Canterbury regions and the number of assaults resulting in hospitalisation in the 17 districts roughly comprising those regions.

2.4.2.1 Seasonal variation in temperature and assaults resulting in hospitalisation.

As for recorded assaults, assaults resulting in hospitalisation were most common in the warmer months (with a peak of 0.21 daily assaults per 100,000 in December), and least common in the colder months—although the nadir was actually in the autumn month of April (0.16 per 100,000 per day).

The effect of seasonal variation in temperature was estimated using a generalised linear mixed model. Seasonal variation in temperature was operationally defined as the seasonal norm (mean) temperature of each of the 365 days of the calendar year within each district, across the entire study period. The response variable was the number of assaults resulting in hospitalisation for each of the 365 days of the calendar year, also across the entire 17-year study period (i.e., the sample size was 365 calendar days*66 districts = 24,090). Population size was controlled, and the intercept was free to vary across districts to account for geographical differences in assault rate. There was little evidence of autocorrelation in these analyses (e.g., average lag 1 r = 0.03 across districts), so error correlations were not included in the model.

The resulting estimate of the effect of seasonal variation in temperature was fairly large, $\hat{\beta}$ = 0.023, 95% CI [0.020, 0.026]. However, when daylength was added as a control variable, the effect of seasonal variation in temperature again seemed to nearly disappear, falling to $\hat{\beta}$ = 0.004, 95% CI [-0.001, 0.009]. Seasonal variation in temperature and daylength were correlated an average r = .82 across districts.

2.4.2.2 Irregular daily variation in temperature and assaults resulting in hospitalisation.

The effect of irregular variation in temperature on assaults resulting in hospitalisation was estimated using a generalised linear mixed model. A Poisson distribution was specified, with population size controlled, and the intercept again free to vary across both districts and years within districts. The response variable was the number of assaults resulting in hospitalisation occurring on a given date within a district, resulting in a sample size of 6209 dates*66 districts = 409,794. Inspection of the residuals indicated that virtually no autocorrelation was present in these analyses, so neither autoregressive terms nor dummy variables for month and weekday were included.

The estimated effect of temperature in this analysis was $\hat{\beta}$ = 0.017, 95% CI [0.013, 0.021], suggesting about 1.7% extra assaults for every 1°C increase in temperature. The potential lagged effects of temperature anomaly on assaults resulting in hospitalisation were again investigated with lags up to 7 days in an alternative model. There was a statistically significant positive effect at lag 1, $\hat{\beta}$ = 0.007, and a significant negative effect at lag 6, $\hat{\beta}$ = -0.008. However, the lagged effects were all small in size and mixed in sign, suggesting little evidence of a substantial lagged effect of temperature. When the humidity anomaly was added to the model, it had no substantial main effect, $\hat{\beta}$ = -0.0007, 95% CI [-0.0017, 0.0003], although there was a positive interaction between temperature and humidity, $\hat{\beta}$ = 0.0006, 95% CI [0.0002, 0.0010], implying a greater effect of temperature at higher humidity levels (or vice versa).

2.4.2.3 Geographical variation in temperature and assaults resulting in hospitalisation.

The mean temperature by district varied from 9.5°C at Queenstown-Lakes to 15.8°C in the Far North District. The correlation between mean temperature and the rate of assaults causing hospitalisation by district was positive, r = .41, 95% CI [.18, .59]. A generalised linear model was also used to estimate the effect of geographical variation in temperature on assaults resulting in hospitalisation by district, as summed across the study period. Using a similar method to the models reported previously, albeit without the need for any random effects, a quasi-Poisson model was estimated. The effect of population size was controlled.

This simple model resulted in a large positive coefficient for temperature, $\hat{\beta}$ = 0.113, 95% CI [0.053, 0.175]. However, a second model was also estimated including demographic controls: the percentage of each district's population that were European, the percentage Māori (European and Māori being the two largest ethnic groups in NZ), the percentage aged 15–39, the percentage 40–64, and the percentage 65+. After these controls, there was no evidence of a positive effect of temperature, $\hat{\beta}$ = -0.029, 95% CI [-0.089, 0.031]. The only demographic control with a statistically significant coefficient was the percentage Māori, $\hat{\beta}$ = 0.022, 95% CI [0.009, 0.035].

We also estimated a third model controlling for regional differences in socioeconomic deprivation, along with the other demographic variables from the second model mentioned above. In this third model, the effect of temperature remained more or less unchanged, $\hat{\beta}$ = -0.026, 95% CI [-0.085, 0.034]. Deprivation level itself had only a small relationship with assault incidence, $\hat{\beta}$ = 0.003, 95% CI [-0.001, 0.007].

2.5 Discussion

Our analysis of the relationship between irregular daily variation in temperature and recorded assault incidence suggested a modest positive effect of temperature, with around 1.5% more assaults for every 1°C increase in temperature. This implies around 3.3% extra assaults for each 1σ temperature increase. This estimate is reasonably similar to the 2.3% per 1σ estimate found in a meta-analysis by Hsiang et al. (2013). Interestingly, the relationship between irregular variation in temperature and assault was very similar in size to that found between temperature and suicide in a recent New Zealand study (1.8% per °C in Williams, Hill, & Spicer, 2015). The estimated effect of temperature on assaults resulting in hospitalisation (i.e., the most serious assaults) was very similar to the effect on recorded assaults. An analysis of the relationship between ENSO and recorded assaults also suggested a similar effect size of a 2% increase in assaults per 1σ SOI increase, demonstrating the plausibility of a link between global meteorological conditions and assault incidence in New Zealand.

Bivariate analyses of seasonal and geographical variation in temperature similarly suggested that assaults were higher during warmer times of the year, and in warmer geographical regions, echoing the findings of previous studies (e.g., Anderson et al., 2000; Hipp et al., 2003). However, the apparent positive effects of these two forms of temperature variation were not robust to controls for daylength (for seasonal variation), or age and ethnicity (for geographical variation). These components of temperature variation are admittedly subject to potential confounding by variables other than those controlled: Other demographic and socioeconomic variables may confound the analysis of geographical variation, while seasonal patterns of human activity not directly caused by temperature may confound the analysis of the effects of seasonal variation in temperature. As such, the estimates reported for the effects of these components of variation may be either upwardly or downwardly biased. The analyses of seasonal and (especially) geographical variation in temperature were also subject to much greater sampling error than the analysis of irregular daily variation: The analysis of geographical variation in temperature and assault was based on a sample size of just 66 districts. As such, sampling error might be to blame for the lack of evidence for a positive effect of temperature in the controlled multivariable models.

This said, despite their limitations, analyses of seasonal and geographical variation do have an advantage over analyses of irregular daily variation in temperature. These analyses allow for some examination of the effects of sustained and predictable differences in temperature as opposed to random day-to-day variations. As such, they may have some value when attempting to draw inferences about the likely effects of sustained climate change. The fact that we did not find strong evidence that seasonal and geographical differences in temperature had a positive effect on assault incidence implies uncertainty about whether climate change will have an effect on assault rates that is adequately approximated by the apparent effects of irregular, day-to-day variation in temperature.

Another source of information bearing on the potential effect of sustained climate change is the analysis of relationships between long-term trends in temperature and violence. Such an analysis requires a longer time series than that examined in this study, although little obvious relationship was apparent between the trends in temperature and assault incidence over the study period (see Figure 4).

2.5.1 The relationship between temperature and assault: Form and mechanism.

The analyses used in our study assumed linear relationships between temperature and (the logarithm of) assault counts. Some previous studies have found evidence of a reduction in assaults with very high temperatures, with an inflection point of around 90°F (32.2°C; Bell & Fusco, 1989; Cohn & Rotton, 2005; Gamble & Hess, 2012). We did not find evidence of such a non-monotonic relationship, although temperatures in New Zealand do not generally reach levels high enough to test this proposition (the highest daily mean temperature observed in the study data was 27.7°C). Non-monotonicity in the relationship between temperature and interpersonal violence is a prediction of one theory that might explain the relationship between these two variables: the negative affect escape model (Bell & Baron, 1976). As mentioned in the introduction, other explanations that have been advanced in the literature include an effect of temperature on opportunity for violence (Cohn & Rotton, 2000), an effect of thermal discomfort on aggression (Anderson et al., 1995), and the hypothesis that higher temperatures increase alcohol consumption. We did not report any tests of these theories in the present study: These theories make only qualitative (and not quantitative) predictions about how temperature will affect violence, and as such, they cannot be used directly to generate quantitative predictions about future changes in interpersonal violence that may occur due to climate change. As such, their practical value when attempting to draw inferences about the effects of climate change may be somewhat limited.

This said, the observation that temperature had a positive relationship with assault incidence even at relatively cold temperatures is inconsistent with the idea that the relationship between temperature and assault occurs because warmer temperatures produce discomfort. Similarly, our analyses of the effect of humidity have some bearing on the plausibility of the theory that assaults rise in hotter temperatures due to thermal discomfort. In the recorded assault data, irregular variation in humidity had a *negative* main effect on assault incidence, and no significant interaction with temperature. These findings are not consistent with the thermal discomfort theory, since higher humidity levels increase thermal discomfort in warm weather. Previous studies have similarly found small or even negative relationships between humidity and assault incidence (Gamble & Hess, 2012; Rotton & Cohn, 2000, 2004). On the other hand, our findings when analysing assaults resulting in hospitalisation were a little more consistent with the thermal discomfort theory: No statistically significant main effect of humidity was found, but there was a significant and positive interaction between temperature and humidity.

2.5.2 How will rising temperatures affect rates of interpersonal violence?

Projections for regional warming in New Zealand were reported by the Ministry for the Environment (2008). Across twelve general circulation models and six emissions scenarios the mean estimate of warming by 2090 was 2.1°C. Perhaps the best estimate of the effect of temperature available in our study can be obtained from the analysis of the effects of irregular daily variation in temperature on recorded assaults, irregular variation being the component of variation least vulnerable to confounding. The resulting coefficient for temperature of 0.015 and its 95% confidence interval of [0.012, 0.017] implies that a 2.1°C increase in temperature might result in between 2.6% and 3.7% more assaults than would otherwise occur. However, across all the models reported in this article—including those analysing the effects of seasonal and geographical variation in temperature—the best estimate of the effect of a 2.1°C increase in temperature ranged widely. The projections vary from a 5.9% decrease in assaults (based on the geographical variation model for hospitalisations with age and ethnicity controlled), to a 26.9% increase (based on the same model without demographic controls). This range of predictions suggests the presence of substantial uncertainty about the effect of temperature on future assaults.

Furthermore, even if climate change does have an increasing *effect* on assault rates, this does not necessarily imply that the actual future assault rate in New Zealand will be higher than today's. Any effect of rising temperatures may be superimposed on other changes in the assault rate, including phenomena such as the long-term decline in human violence in recent centuries and millennia (Pinker, 2011).

On the other hand, climate change may affect interpersonal violence via mechanisms that may not have observable effects in the data studied here, meaning that there is a risk that the models estimated in this study could underestimate the true effect of climate change on assault incidence. Anderson and DeLisi (2011) suggest, for example, that food scarcity and environmental disasters may increase violent crime rates. Agnew (2011) argues that climate change will increase crime via a number of pathways, including impacts on social strain, support and control structures, and social conflict arising due to competition over scarce resources, migrations, and poverty. As well as potential effects on interpersonal violence, climate change may also impact the probability of conflict between groups and states by compromising economic security, culture, and identity, altering migration flows, and harming the ability of states to provide important infrastructure and public services (Adger et al., 2014). Indeed, there is some empirical evidence that temperature increases heighten the probability of intergroup conflict (e.g., Hsiang et al., 2013), although this hypothesis is still contested in the literature (Adger et al., 2014; Salehyan, 2008). Given that global warming of the magnitude likely over the following century has not been observed since long before humans started to collect assault statistics, producing accurate predictions about how the incidence of violence will change in a warming world remains a challenging task. Despite the positive relationship between temperature and the incidence of interpersonal violence, great uncertainty remains about how the incidence of violence will change in a warming world.

2.6 Conclusions and Commentary

Each empirical study reported in this thesis is followed by a conclusions and commentary subsection in which I briefly integrate the findings of the study into the thesis as a whole. In Study One, I found that temperatures had a positive bivariate relationship with the incidence of assault. This bivariate relationship held regardless of whether the component of temperature variation examined was irregular, seasonal, or geographical. Our findings with respect to the effect of irregular variation in temperature were consistent with those of another New Zealand study reporting a modest positive effect of irregular variation[3] in temperature on the incidence of violent crime in the period 2000–2008 (Horrocks & Menclova, 2011). There seems to be a reasonable case for interpreting the bivariate relationship between irregular variation in temperature and assault as causal. Irregular daily variation in temperature is by definition correlated neither with regional differences in demographic and economic variables, nor with calendar variables such as time of year. As such, irregular daily variation in temperature seems to be the component of variation whose apparent effects on assault rates (and other behavioural variables) are least vulnerable to confounding.

On the other hand, the apparent effects of seasonal and geographical variation in temperature are more vulnerable to confounding. It is easy to envisage variables that

[3] Horrocks and Menclova (2011) used daily data from 43 police districts. Their use of control variables for district and month of year meant that their analysis implicitly focused on the effects of irregular variation in temperature (although they did not use this terminology). Their "violent crime" dependent variable included crimes of assault, intimidation, homicide, property damage, and property abuse.

might be correlated with geographical or seasonal variation in temperature, that are not themselves affected *by* temperature, and that may affect assault rates (thus having the formal characteristics of a confound; see McNamee, 2003). The application of statistical controls for some obvious examples resulted in the apparent effects of seasonal and geographical variation in temperature becoming non-significant. This does not necessarily mean that these components of variation had no effect. Problems such as the limited sample size (for geographical variation) and the strong correlation between temperature and the potential confound of daylength (for seasonal variation) made it harder to accurately estimate the size of the effects of these components of variation in temperature. Nevertheless, the fact that the estimated effects of seasonal and geographical variation in temperature were not overtly consistent with that of irregular variation in temperature meant that this study could not rule out the possibility that sustained changes in climate may have very different effects from those of irregular day-to-day variation in temperature. The study ended with some predictions of the effects of anthropogenic climate change on assault rates in New Zealand: The best estimate (based on the analysis of irregular variation in temperature and recorded assaults) was of a small positive effect of climate change on assault incidence. However, there was a great deal of uncertainty surrounding this estimate.

Overall, this study provided some tentative evidence that warmer temperatures in New Zealand may result in higher rates of interpersonal violence. But would this apparent effect also hold for acts of violence committed against the self, such as acts of suicide? This is the topic of the next empirical study.

3 Study Two: Will Climate Change Increase or Decrease Suicide Rates? The Differing Effects of Geographical, Seasonal, and Irregular Variation in Temperature on Suicide Incidence

The following article has been published in *Climatic Change*. The version shown is the accepted manuscript after peer review, with minor formatting changes for consistency with the rest of this thesis, and some typographical errors corrected. An electronic supplementary materials document that was provided along with the main text is also reproduced in Appendix B (section 10), and contains additional technical details about the study's method and results. References for this article are provided within the consolidated reference list. The final publication is available at Springer via the citation below.

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Will climate change increase or decrease suicide rates? The differing effects of geographical, seasonal, and irregular variation in temperature on suicide incidence. *Climatic Change*, *130*(4), 519–528. http://doi.org/10.1007/s10584-015-1371-9

3.1 Abstract

The effect of environmental temperature on suicide risk is an important issue given the increase in global temperatures expected over the following century. Previous research has produced conflicting findings: Studies concerned with temporal variation in temperature and suicide have tended to find a positive relationship, while those concerned with geographical variation in temperature and suicide have tended to find a negative relationship. In this study, we aimed firstly to estimate the relationship between suicide incidence and three components of variation in temperature: irregular, seasonal, and geographical. Secondly, we aimed to critically examine what this information can (and cannot) tell us about the likely effects of anthropogenic climate change on suicide rates. Suicide data from New Zealand for the period 1988 to 2007 were collated according to date of death and district and compared with temperature data from the same period. Using generalised linear mixed models, we found that irregular variation in temperature was positively related to suicide incidence, with about 1.8% more suicides for every 1°C increase in temperature. On the other hand, seasonal variation in temperature had virtually no linear relationship with suicide incidence, and

when controlling for demographic differences, geographical variation in temperature was *negatively* related to suicide incidence. We conclude that differences in both the sign and the direction of the effects of different forms of variation in temperature mean that it is very difficult to predict how climate change will affect risk of suicide.

3.2 Introduction

There is a substantial literature concerned with the relationship between climatic factors and the incidence of suicide (for reviews see Deisenhammer, 2003; Dixon & Kalkstein, 2009), with much of this literature concerned with the effect of temperature on suicide risk. An apparent relationship between temperature and suicide leads naturally to questions about how anthropogenic climate change will affect suicide rates. At least one recent study has suggested that global warming will increase the risk of suicide (Preti et al., 2007). This said, the reviews by Deisenhammer (2003) and Dixon and Kalkstein (2009) both noted that existing findings concerned with the effect of temperature on suicide are contradictory: A number of studies have found a positive relationship between temperature and suicide incidence (Likhvar, Honda, & Ono, 2011; Maes, Meyer, Thompson, Peeters, & Cosyns, 1994; Müller et al., 2011; Page et al., 2007; A. C. Yang, Tsai, & Huang, 2011; Yan, 2000), while others have found a negative relationship (Lester, 1999; Robbins et al., 1972; Rotton, 1986; Souêtre, Wehr, Douillet, & Darcourt, 1990).

This apparent paradox can be at least partially resolved by noting an important difference between the studies finding a negative relationship and those finding a positive relationship. Specifically, studies finding a negative effect tend to be those estimating the effects of *geographical* variation in temperature (i.e., comparing suicide rates across areas with different mean temperatures), whereas those finding a positive effect tend to be concerned with the effects of *temporal* variation in temperature (i.e., comparing different time periods within the same geographical area).

In an example of the geographical comparison approach, Lester (1999) compared 62 nations and found that temperature was negatively correlated with suicide rates, $r = -.59$. This finding was similar to that of an earlier international study finding a correlation of -.58 between suicide rate in 1965 and a 6-point scale of climate warmth (Robbins et al., 1972). Another study of 48 countries by Rotton (1986) found a negative correlation of similar magnitude between mean January temperatures and suicide rate, r

= -.58. A study of 19 French regions also found that regional mean temperature was negatively correlated with suicide rates, r = -.5 in both 1975 and 1983 (Souêtre et al., 1990). Similarly, an Italian study found a correlation of -.45 between annual mean minimum temperature and suicide rate across 17 towns, although this relationship was not quite statistically significant (Preti, 1998).

On the other hand, studies of the effects of temporal variation in temperature have tended to find a positive relationship between temperature and suicide rates. Temporal variation in temperature can be broken down into further sub-components, including seasonal variation and irregular (random) variation. Existing studies suggest that the positive effect of temporal variation seems to hold across these two components. In an example of a study of irregular daily variation in temperature, Page et al. (2007) found a positive relationship between mean daily temperature and suicide risk in Britain, albeit only above a certain threshold: Every 1°C increase in mean daily temperature above 18°C was associated with a 3.8% increase in suicides, when controlling for month. Ajdacic-Gross et al. (2007) likewise found a significant positive association between monthly temperature and both male and female monthly suicides in Switzerland over 1881–2000, after fitting an autoregressive integrated moving average model to each variable to remove seasonal, trend and autocorrelative patterns. Studies of the effects of irregular daily variation in temperature in Mittelfranken, Germany (Müller et al., 2011) and Japan (Likhvar et al., 2011) likewise found positive effects on suicide incidence, as did a study using data from both Toronto and Jackson, Mississippi (Dixon et al., 2014).

A seasonal pattern in suicide deaths is well established, with a peak in spring (see Ajdacic-Gross, Bopp, Ring, Gutzwiller, & Rossler, 2010 for a review). While most studies of seasonality in suicide have examined the univariate distribution of suicides, a small number of studies have specifically assessed the relationship between seasonal variation in temperature and suicide. Maes et al. (1994) showed that violent suicide rate, sunshine duration and ambient temperature shared a common annual rhythm in Belgium. Similarly, Preti (1997) found a positive correlation of 0.67 between mean temperature (across the 12 months of the calendar year) in Italy and the mean number of suicides. This said, the fact that suicides tend to peak in the spring rather than summer suggests that a linear effect of temperature is not the major factor driving seasonality in suicides.

Admittedly, not all studies on the topic fit the pattern of a positive effect of temporal variation in temperature, but a negative effect of geographical variation. A study in Queensland found a positive relationship between geographical variation in temperature and suicide incidence (Qi, Tong, & Hu, 2009). Similarly, not all temporal studies have found positive effects of temperature: A study of suicides in North Carolina found no significant effect of temperature on suicide incidence (Zung & Green, 1974), nor did a study of a township in Greenland (O. Grove & Lynge, 1979). Overall, however, the pattern of a negative effect of geographical variation in temperature and a positive effect of temporal variation in temperature seems quite well substantiated in most locations.

No general review of the effect of climatic factors other than temperature on suicide incidence is provided here for brevity's sake. However, one climatic factor is of particular relevance: solar radiation. Global solar radiation (Müller et al., 2011; Ruuhela, Hiltunen, Venäläinen, Pirinen, & Partonen, 2008) is positively related to suicide incidence, as is number of sunshine hours (Vyssoki et al., 2012). Solar radiation and temperature are also related. Therefore, solar radiation is a potential confound of the temperature-suicide relationship. As such, it can serve a useful role as a control in multivariable analyses.

3.2.1 Explanations of the effects of temperature on suicide.

The mechanism linking temperature and suicide risk is not well understood. A number of theories seek to explain seasonality in suicides, including greater perceived opportunity of some suicide methods in warmer weather (Ajdacic-Gross et al., 2010), seasonal changes in serotonergic function (Maes et al., 1995), and Durkheim's (1897) theory that the spring suicide peak is caused by a greater intensity of social life. A recently proposed theory relates to the observation that brown adipose tissue is activated in cold temperatures to produce non-shivering thermogenesis (Holopainen, Helama, Björkenstam, & Partonen, 2013; Holopainen, Helama, & Partonen, 2014). Holopainen et al. suggest that, in spring, rapidly increasing temperatures in the absence of an inhibitory long photoperiod may result in relative over-activation of brown adipose tissue, a condition believed to cause symptoms similar to those of depression. This said, a suicidogenic effect of over-activation of brown adipose tissue relative to ambient temperatures cannot explain why geographical variation in temperature and suicide are negatively related. In general, no complete single explanation for the varying relationships between suicide and different components of variation in temperature exists.

3.2.2 Consideration of global climate change.

Research studies concerned with the effects of temperature on intrapersonal violence have thus far given little direct focus to the problem or implications of climate change, aside from passing mentions of the issue (e.g., Page et al., 2007; Törő et al., 2009; Y. Kim, Kim, & Kim, 2011). One exception is a study using national Italian data (Preti et al., 2007). Preti et al. claimed to have found evidence for a link between global warming and an enhanced risk of suicide. However, Preti et al. did not consider the issue of the conflicting effects of geographical and temporal variation in temperature.

3.2.3 Aims of the current study.

This study aimed to estimate the relationship between the incidence of suicide and three components of variation in temperature in New Zealand: irregular variation, seasonal variation, and geographical variation. We also aimed to critically examine what the resulting information can (and cannot) tell us about the likely effects of future climate change on suicide rates.

3.3 Methods

3.3.1 Suicide data.

A listing of nationwide suicide deaths from New Zealand was obtained from the Ministry of Health for the period 1 January 1988 to 31 December 2007. The date of death, gender, and territorial local authority (i.e., district) of residence of the deceased were included. Deaths due to late effects of self-harm were excluded. These data were then collated into a count of suicides on each date and for each of the 67 current districts of New Zealand.

3.3.2 Meteorological data.

Meteorological data were obtained from the National Institute of Water and Atmospheric Research’s virtual climate network, which covers New Zealand on a regular 5km grid. The virtual station closest to the town centre of the largest town or city within each district was used to represent the given district. Daily mean temperatures were calculated as the mean of the daily minimum and maximum temperatures. Seasonal norm temperatures were obtained by calculating the mean temperature (across the entire period of 20 years) for each day of the 365 days of the calendar year in each district, and then subtracting the overall mean temperature for that district. In order to operationally define irregular variation in temperatures, the temperature anomaly for every date-district combination was defined as the difference between the observed temperature on a given date, and the average temperature for that location and day of the year.

3.3.3 Population data.

Annual population estimates by district were obtained from Statistics New Zealand. Demographic (age and ethnicity) information for each district as at the 1996, 2001 and 2006 censuses was obtained from the Statistics New Zealand website (Statistics New Zealand, 2013).

3.3.4 Data analysis.

Data analysis was completed in R version 3.0.2 (R Core Team, 2013), with the lme4 package version 1.0-5 (Bates, Maechler, Bolker, & Walker, 2013) used for generalised linear mixed models. A Poisson distribution with a log link was used for mixed models. The effect of population size was controlled in all substantive analyses, following a log transformation to allow population size to have an additive effect. Spline analysis was completed in the crs package (Nie & Racine, 2012), with automatic selection of the number of segments and polynomial form via Kullback-Leibler cross-validation. Visually-weighted regression (Hsiang, 2013) was completed using R code by Schönbrodt (2012), with the span for the loess smoother selected using the bias-corrected Akaike information criterion, as implemented in the fANCOVA package (X.-F. Wang, 2010).

For further information about methods and data sources please refer to the Electronic Supplementary Materials.

3.4 Results

In the 20 years surveyed, 9984 suicides were recorded across the 67 districts of New Zealand (excluding a further 47 suicides with a district of “overseas/other” listed). A clear majority of these suicides (78%) were by males. The mean national suicide rate over the study period (1988 to 2007) was 13.3 suicides per 100,000 per annum. Over the same period, the mean temperature across all 67 districts and 7305 days of the study period was 12.7°C (*SD* = 4.3°C).

The first analysis of the effect of temperature was a simple one in which variation in temperature was not isolated into geographical, seasonal, and irregular components. This was completed by specifying a Poisson generalised linear mixed model, with the number of suicides for each date and district in the dataset as the response, and the only predictors being mean temperature and the logarithm of population size. This model produced a positive but very small coefficient for temperature of $\hat{\beta}$ = 0.004, with the 95% confidence interval including zero, CI [-0.001, 0.009]. While this finding might seem to suggest no or little effect of temperature on suicide incidence, the sections that follow show how different types of variation in temperature appear to have quite different effects on the incidence of suicide.

3.4.1 Effects of irregular variation in temperature.

The effect of irregular variation in temperature was examined by entering temperature anomalies as a predictor into a Poisson generalised linear mixed model, with population controlled, and the intercept free to vary across districts. The number of suicides occurring for each date-district combination was the response variable. In this model, the coefficient of 0.018, 95% CI [0.009, 0.027] for irregular variation in temperature suggested that every 1°C temperature increase was associated with approximately 1.8% more suicides.

The estimated effect of temperature was robust to alternative modelling choices, with the point estimate remaining at 0.018 (within rounding) given alternative strategies such as adding a control for radiation and using a negative binomial rather than Poisson model. An alternative non-linear analysis was also completed, estimating the relationship between nationally averaged temperature anomaly and national suicide rate by date using loess smoothing (see Figure 5). As is visible, the relationship between temperature anomaly and suicide incidence is fairly well approximated by a linear form.

Indeed, a spline analysis of suicide rate per capita (across all dates and districts) regressed on temperature anomaly also suggested that a simple linear model was the best fit, justifying the linear model reported. Figure 5 also shows that while the effect of irregular variation temperature can be estimated with reasonable precision for temperature anomalies between roughly -3°C and 3°C, the effects of temperatures outside this range are subject to much greater uncertainty.

We also attempted to determine whether irregular temperature had delayed effects on suicide incidence by calculating lagged temperature terms. We added lagged temperature terms up to a lag of seven days to the generalised linear mixed model mentioned above. In this model, the contemporaneous effect of temperature remained similar, but the lagged effects were all small and not statistically significant.

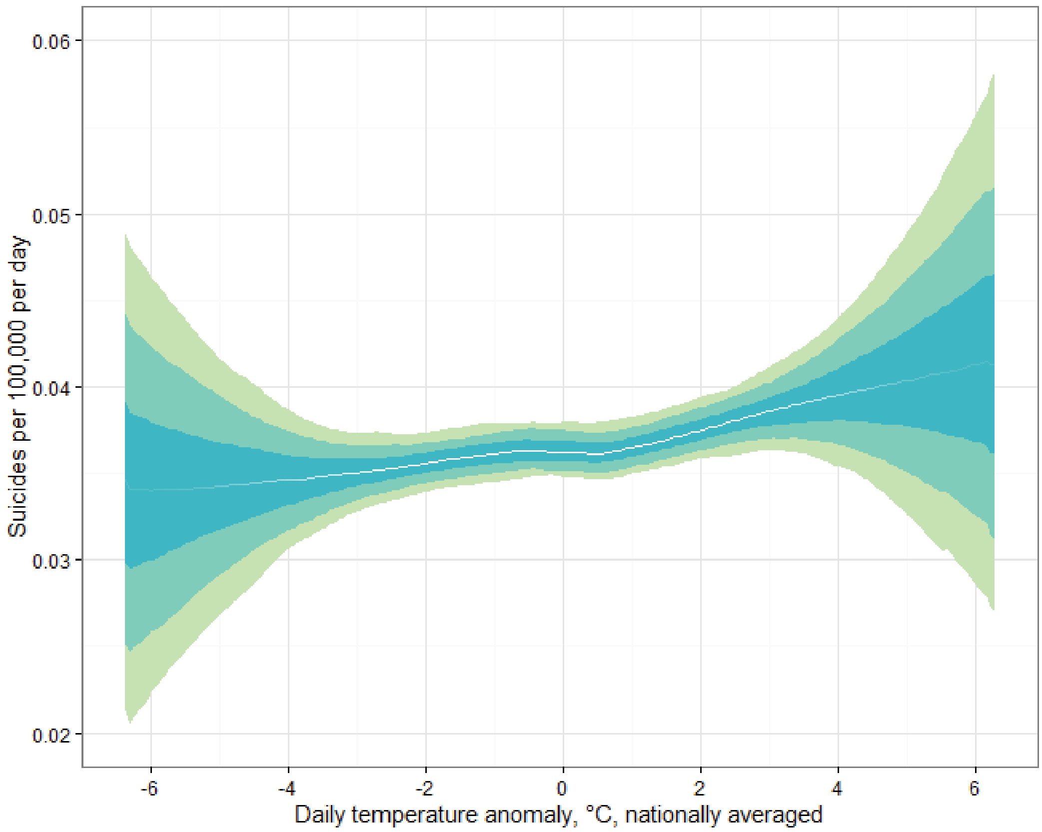

Figure 5. Visually-weighted non-linear regression analysis of the effects of irregular variation in temperature. Model estimated using loess smoothing (degree = 2, span = 0.9). The line of best fit is in white. The shaded bands indicate the regions within 1, 2, and 3 standard errors of the estimate.

3.4.2 Effects of seasonal variation in temperature.

Mean daily suicides rates and mean temperatures by month, across all districts, are displayed in Figure 6. The seasonal pattern is qualitatively in accordance with the usual findings in the field: The number of suicides peaks in the Southern Hemisphere spring, with a trough in winter. However, the magnitude of the seasonal variation is small.

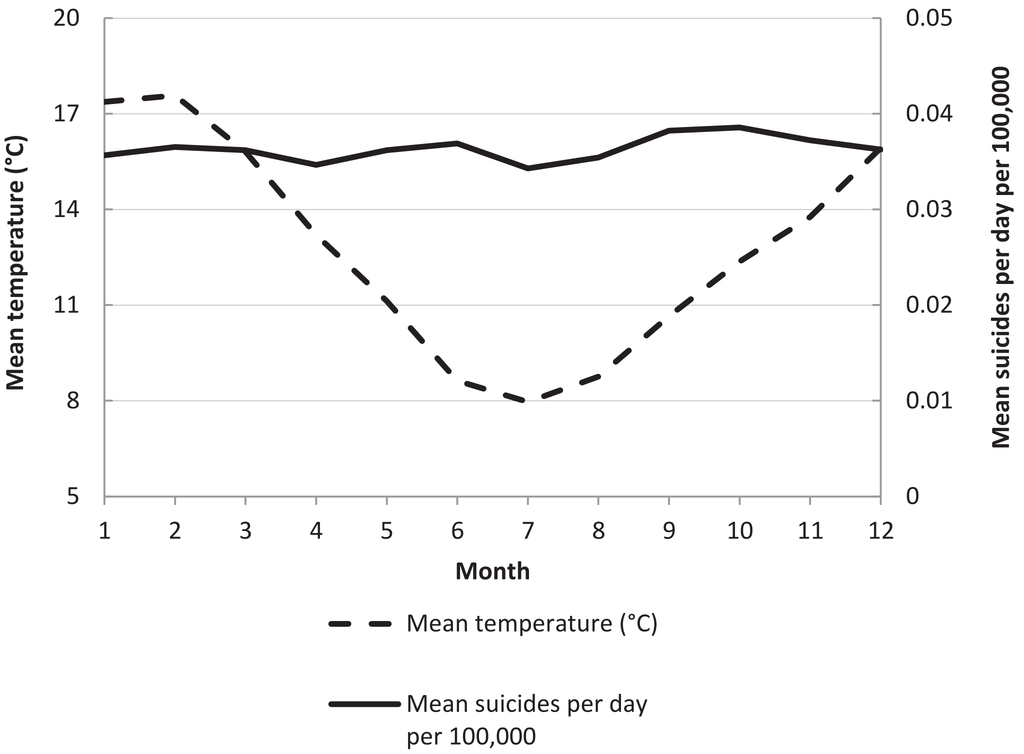

Figure 6. Mean daily suicides per 100,000 and temperature by month.

The impression of a weak relationship between seasonal variation in temperature and suicide incidence was corroborated by a generalised linear mixed model in which the total number of suicides summed over the 20 occurrences of each of the 365 days of the calendar year (across the entire study period) for each district was the response variable, while seasonal norm temperatures and population size were predictors. The intercept was specified as a random effect across districts. The resulting estimated effect of seasonal variation in temperature[4] was tiny, $\hat{\beta} = 4.63 \times 10^{-4}$, $p = .881$, 95% CI [-0.006, 0.007].

[4] In the original article (Williams, Hill, & Spicer, 2015), this sentence incorrectly stated that the "resulting estimated effect of *monthly temperature* [emphasis added] was tiny" (p. 524). The unit of analysis here was days of the calendar year, not months. An erratum has been submitted to the journal.

3.4.3 Geographical variation in temperature.

A scatter plot of mean temperature by geographical location versus suicide rate is displayed in Figure 7. No clear relationship is evident. The correlation between geographical mean temperature and suicide rate was indeed close to zero, $r(65) = .094$, 95% CI [-.150, .326]. Similarly, a negative binomial model with a control for population suggested a small effect of temperature, $\hat{\beta} = .005$, with a wide 95% confidence interval that spanned zero, [-0.022, 0.032].

However, an alternative analysis was also completed with controls for age and ethnicity (European percentage of the population, Māori percentage of the population, percentage aged 15–39, percentage aged 40–64, and percentage aged 65+). In this analysis, the estimated effect of temperature became negative, $\hat{\beta} = -0.034$, 95% CI [-0.067, -0.001]. Further adding a control for radiation resulted in the estimate remaining negative but falling outside significance, $\hat{\beta} = -0.026$, 95% CI [-0.065, 0.012], with the effect of radiation likewise being non-significant, $\hat{\beta} = -0.020$, 95% CI [-0.073, 0.033].

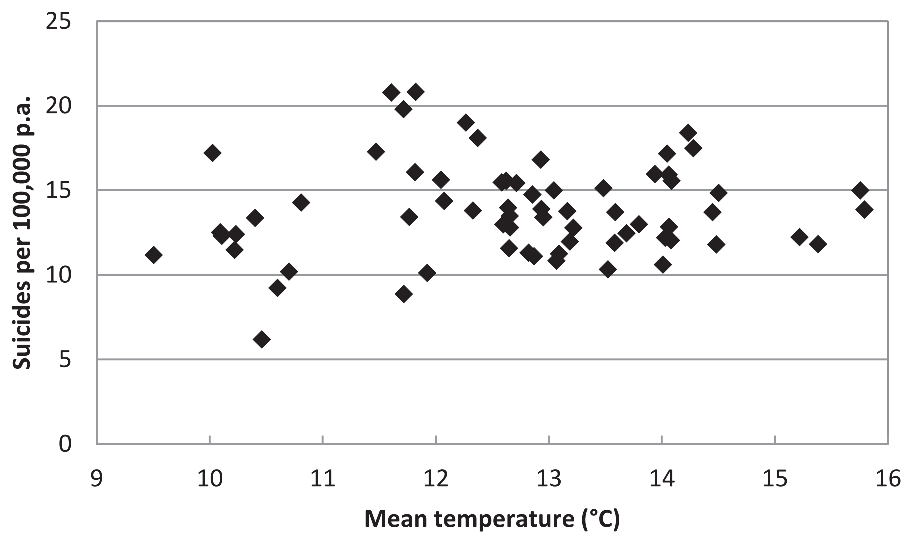

Figure 7. Geographical variation in temperature versus suicide rate.

3.5 Discussion

In this study, irregular variation in temperature had a positive relationship with suicide incidence, with approximately 1.8% more suicides for every 1°C increase in temperature. The size of this estimate was fairly consistent with those of previous studies (e.g., Deisenhammer, 2003; Y. Kim et al., 2011). We did not find evidence of the non-linear relationship between temperature and suicide incidence reported by Page et al. (2007). We also found no evidence for any substantial lagged effect of irregular variation in temperature, a finding similar to that of Likhvar et al. (2011) and Y. Kim et al. (2011).

On the basis of the apparent positive effect of irregular variation in temperature, it would be tempting to conclude that global warming will increase the incidence of suicides. However, this conclusion is contradicted by our analyses of seasonal and geographical variation in temperature and suicide. Seasonal variation in suicide deaths did roughly follow the pattern generally found in the literature of a peak in spring and a trough in winter (Chew & McCleary, 1995), but in accordance with a previous study in New Zealand (Yip, Chao, & Ho, 1998), the magnitude of seasonal variation in suicide incidence was very small. Furthermore, seasonal variation in temperature had a relationship with suicide incidence that was close to zero: That is, hotter times of the year were not consistently associated with higher suicide risk.

The relationship between geographical variation in temperature and suicide incidence differed even more greatly from that of irregular variation in temperature. While there was little evidence of any relationship at a bivariate level, controlling for age and ethnic differences across regions resulted in the relationship becoming negative: Warmer areas had *lower* suicide rates. This finding was in accordance with prior studies finding a negative relationship between geographical variation in temperature and suicide rates (Lester, 1999; Rotton, 1986; Souêtre et al., 1990). The estimated effect became non-significant when controlling for radiation, although the effect of radiation was likewise not significant. As such, it is difficult to confirm whether the apparent relationship between geographical variation in temperature and suicide is due to a unique effect of temperature, or just to differences in sunlight exposure. Different types of variation in temperature therefore appeared to have quite different effects on suicide incidence in this study. One potential explanation is that one or more of the analyses were confounded. The effects of irregular variation in temperature seem least likely to be confounded, given that short-term day-to-day temporal variation in temperature is random, relatively unpredictable, and not strongly influenced by or related to human factors (e.g., economic, demographic, or social variables). Furthermore, the variables that do explain irregular variation in temperature are meteorological in nature, and can be controlled for statistically (solar radiation being a prime example). On the other hand, analyses of geographical variation are much more susceptible to confounding by demographic, economic, and social variables that differ across regions and affect suicide rates. In the current study, we controlled for two plausible confounds (age and ethnicity), but it is difficult to rule out the possibility of other confounds. Statistical control for other potential confounds (e.g., economic production, differences in divorce rates) was difficult to apply given the lack of information available about inter-district differences in many social and demographic variables. The lack of such controls in the geographical analysis is a limitation of this study. The effect of seasonal variation in temperature on suicide incidence is likewise somewhat susceptible to confounding by seasonal and calendar-related cycles in human activity that are not caused by temperature.

Counteracting the potential confounding problem relating to analyses of seasonal and (especially) geographical variation is the fact that these analyses can potentially provide information about how humans adapt to long-term, sustained climatic differences. Indeed, the negative relationship between geographical variation in temperature and suicide incidence hints at the presence of adaptation mechanisms to warmer temperatures that inhibit suicide risk in the long term. At the very least, this finding means that we should be cautious about assuming that the positive effect of irregular variation in temperature on suicide incidence implies that climate change will increase suicide risk. Until the reasons for the apparently conflicting effect of geographical variation in temperature can be identified, such an inference does not seem well justified.

What would be particularly useful is a theoretical account that can explain all of the various empirical regularities established with respect to the relationship between temperature and suicide incidence: the negative relationship between geographical variation in temperature and suicide, the positive relationship between irregular variation in temperature and suicide, and the established seasonal pattern variation in suicide rates (i.e., a springtime peak and winter nadir). Existing explanations generally attempt only to explain one or two of those regularities; in fact most theoretical accounts focus primarily on explaining seasonality in suicides. Seasonal variation in suicide has been explained as being caused by differences in perceived availability of some suicide methods (Ajdacic-Gross et al., 2010), by over-activation of brown adipose tissue in spring (Holopainen et al., 2013), or by changes in serotonergic functioning (Maes et al., 1995). The development of a coherent theoretical account explaining the relationships observed between different types of variation in temperature and suicide incidence could provide guidance when attempting to generate predictions about the impacts of climate change, and is an important avenue for future research. We did not attempt to produce or test such a framework, which was a limitation of this study. In the absence of a coherent theoretical account, differences in both the size and the direction of the effects of different types of variation in temperature means that it remains difficult to predict both the direction and the size of the future effect of climate change on suicide incidence.

3.6 Acknowledgements

We thank Jane Perrott and Chris Lewis at the Ministry of Health for providing suicide mortality data, Richard Speirs at Statistics New Zealand for providing population data, and two anonymous reviewers for helpful feedback.

3.7 Conclusions and Commentary

Study One suggested that irregular variation in temperature was associated with an increased incidence of assault, with approximately 1.5% extra assaults for every 1°C increase in temperature. Remarkably, Study Two indicated a very similar relationship between irregular variation in temperature and suicide: around 1.8% extra suicides for every 1°C increase in temperature.

In Study One, the bivariate relationship between temperature and the incidence of assault was positive, regardless of the component of variation examined (albeit that there was no strong evidence of positive effects of geographical and seasonal variation in temperature on assault, once some obvious potential confounding variables[5] were controlled for). In Study Two, however, the empirical picture was even more complex.

---

[5] In the seasonal analyses, a potential confounding variable of particular concern was exposure to sunlight, exposure to sunlight being a variable that affects temperature and that might itself directly affect the incidence of violence. This variable was controlled for in some form in all three studies. In Study One (assaults), exposure to sunlight was operationalised simply as daylength, whereas in Studies Two (suicides) and Three (self-harm) it was operationalised as a closely related variable: quantity of solar radiation. The operationalisation as daylength in Study One made sense given that longer days might affect assault rates most plausibly simply by affording more opportunity for interaction between potential victims and offenders. On the other hand, for the studies of suicide and self-harm, there was a concern that there might be a direct biological effect of exposure to sunlight on the risk of intrapersonal violence (see Müller et al., 2011; Ruuhela, Hiltunen, Venäläinen, Pirinen, & Partonen, 2008). A direct control for radiation was therefore applied in Studies Two and Three.

There was little evidence of a relationship between seasonal variation in temperature and suicide incidence even at a bivariate level, despite the large body of research demonstrating seasonality in suicides (see Ajdacic-Gross et al., 2010). Furthermore, when a selection of plausible confounding variables was controlled for, there was some evidence of a *negative* effect of geographical variation in temperature on suicide risk. This latter finding was bolstered by the presence of previous studies that have found higher suicide rates in colder regions.

Study One concluded with a tentative prediction of the number of extra assaults that might arise as a result of a plausible estimate of the warming in New Zealand likely over the 21st century (2.1°C). The presence of a great deal of uncertainty surrounding this estimate was stressed in Study One. However, in Study Two there seemed to be even less of a basis to provide any kind of quantitative prediction. Specifically, while the estimated effect of irregular variation in temperature was indeed positive, the fact that the estimated effect of geographical variation in temperature was negative implied a great deal of uncertainty not just about the size but even the direction of the likely effect of rising temperatures. If the apparent negative effect of geographical variation in temperature does actually represent how humans react to sustained differences in climate, a warmer world might be one with a lower suicide rate. Given the presence of substantial uncertainty with respect to both the size and the direction of the effect of temperature on suicides, quantitative predictions were not attempted in this case. This reflects the great difficulty in predicting how human behaviour will respond to climate change. This is a topic that will be returned to in the methodological critique section of this thesis.

The analysis of suicides presented in Study Two allowed for some investigation of the relationship between temperature and intrapersonal violence. However, an important feature of this investigation was that its focus on suicides meant that only the most serious acts of intrapersonal violence were studied. On a pragmatic level, this meant that while the dataset studied contained suicide counts for a great number of dates and locations, the actual frequency of the behaviour of interest—suicides—was much smaller than for the analysis of assaults in Study One. Despite the restriction to just three regions of New Zealand, Study One involved the investigation of over 260,000 recorded assaults, whereas fewer than 10,000 suicides were analysed in Study Two.

While neither case can be described as a small sample by any stretch of the imagination, it also seemed valuable to study the relationship between temperature and a somewhat broader domain of intrapersonal violence. Specifically, in the following study (Study Three) I investigated the relationship between temperature and the incidence of acts of self-harm resulting in hospitalisation. A possible effect of temperature increases on the incidence of such acts of intrapersonal violence is interesting and potentially valuable for its own sake: Acts of self-harm serious enough to result in hospitalisation can obviously have very serious consequences for those involved. Furthermore, by studying a broader category of intrapersonal violence, it was possible to obtain an even larger sample of behaviour, and produce estimates with greater precision than was the case for Study Two's investigation of suicides. Study Three is reported next.

4 Study Three: Do Hotter Temperatures Increase the Incidence of Self-Harm Hospitalisations?

The following article has been published (as an advance online publication) by *Psychology, Health, and Medicine*. The version shown is the accepted manuscript, with minor formatting changes for consistency with the rest of this thesis, and some typographical errors corrected. An electronic supplementary materials document that was provided along with the main text is provided in Appendix C (section 11), and contains additional technical details about the study's method and results. References for this article are provided within the consolidated reference list at the end of this thesis. A reference to the final version of the article is shown below.

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Do hotter temperatures increase the incidence of self-harm hospitalisations? *Psychology, Health & Medicine*. Advance online publication. http://doi.org/10.1080/13548506.2015.1028945

4.1 Abstract

A relationship between air temperature and the incidence of suicide has been established in a number of previous studies. Interestingly, the relationship between geographical variation in temperature and suicide incidence has generally been found to be negative, while the relationship between temporal variation in temperature and suicide incidence has generally been found to be positive. It is less clear, however, how temperature relates to the incidence of self-harm. This topic is of particular importance given the presence of ongoing global warming. This study investigated the relationship between temperature and the incidence of self-harm resulting in hospitalisation in New Zealand. Self-harm hospitalisations by date and district for 1993–2009 were obtained from the Ministry of Health. Meteorological data was obtained from NIWA. Generalised linear mixed models were used to estimate the effects of three different components of variation in temperature: geographical, seasonal, and irregular. Irregular (random) daily variation in temperature had a modest positive relationship with the incidence of acts of self-harm resulting in hospitalisation, with about 0.7% extra incidents for every 1°C increase in temperature. However, there was no strong evidence for a positive effect of either seasonal or geographical variation in temperature. We conclude that temperature

does appear to bear some relation to the incidence of self-harm, with irregular daily variation in temperature having a positive effect. However, inconsistencies in the effects of different components of variation in temperature make it challenging to accurately predict how global warming will influence the incidence of self-harm.

4.2 Introduction

The effect of temperature on the incidence of fatal self-harm (i.e., suicide) has been studied since at least the late 1800s (Morselli, 1882). Interestingly, studies comparing different geographical regions have tended to find that colder areas have higher suicide rates (Lester, 1999; Rotton, 1986; Souêtre et al., 1990), while studies investigating the effects of temporal variation in temperature within individual regions have tended to find positive effects (Ajdacic-Gross et al., 2007; Deisenhammer, Kemmler, & Parson, 2003; Helama et al., 2013; Y. Kim et al., 2011; Page et al., 2007). What has been studied less is how temperature relates to the incidence of acts of self-harm more generally, including *non-fatal* self-harm. The issue of the effects of temperature on self-harm is of particular importance given the presence of ongoing anthropogenic global warming (IPCC, 2013a).

The small number of studies that have attempted to investigate the relationship between temperature and the incidence of self-harm have tended to focus on the effects of temporal variation in temperature. Furthermore, the focus has tended to be on seasonal variation. For example, a study of admissions to an emergency department in Turkey (Doganay et al., 2003) found a strong correlation, $r = .81$, between the average number of suicide attempts in each month of the calendar year and temperature. An Italian study of suicide attempts for 1974–1994 also found a strong and positive (but not significant) correlation between mean temperature and the number of suicide attempts across the 12 months of the calendar year, $r = .49$ (Preti, 1997). These findings are consistent with studies of the seasonal distribution of self-harm, which tend to find peaks in the warmer months of spring (Jessen et al., 1999; Jessen, Steffensen, & Jensen, 1998; Rock & Hallmayer, 2008) or summer (Masterton, 1991).

A study of 12,379 parasuicide admissions in Oxford for 1976–1989 (Barker, Hawton, Fagg, & Jennison, 1994) took the alternative approach of analysing data at the daily level. This meant that their study estimated the combined effects of both seasonal and *irregular* (random) variation in temperature. A significant but small correlation between parasuicide admissions and maximum temperature was found for women, r = .04, but not men. A study of hospital admissions in Helsinki attempted to directly estimate the effect of irregular variation in temperature on suicide attempts by calculating the difference between daily observed temperatures and seasonal norms (Hiltunen et al., 2012). The resulting estimated effect of temperature was small and not statistically significant.

In general, the effect of irregular daily variation in temperature on the incidence of self-harm is as yet not well established. Similarly, we were unable to find any study that specifically assessed the effect of geographical variation in temperature on the incidence of self-harm. These two components of variation in temperature are of particular interest for different reasons. The effect of irregular variation in temperature is a valuable object of study because, due to its random and unpredictable nature, irregular daily variation is less likely to be subject to important confounds than other components of temperature variation. On the other hand, while the effect of geographical variation in temperature may be difficult to separate from demographic or cultural differences between populations living in different areas, investigation of the relationship between geographical variation in temperature and self-harm is still valuable: It may help to inform inferences about the effects of long-term, *sustained* differences in climate, such as those occurring due to anthropogenic climate change.

The current study therefore aims to specifically estimate the effects geographical, seasonal, and irregular variation in temperature have on the incidence of acts of self-harm resulting in hospitalisation in New Zealand. Furthermore, this study aims to discuss what this information can (and cannot) tell us about how the incidence of self-harm is likely to be affected by anthropogenic climate change.

4.3 Methods

4.3.1 Description of setting.

New Zealand is an island nation with approximately 4.5 million inhabitants (MacPherson, 2014). Its climate ranges from warm and subtropical in the north to cool and temperate in the south, with severe alpine conditions in some mountainous regions (Mackintosh, 2001). The most common ethnicities are European (74% in 2013), Māori (15%), Asian (12%), and Pacific peoples (7%) (Statistics New Zealand, 2014).

4.3.2 Hospitalisations data.

The New Zealand Ministry of Health provided nationwide data listing public hospital discharges with a discharge date between 1 January 1993 and 31 December 2009 for which at least one of the causes of hospitalisation was self-harm (ICD-9[6] codes E950–E958). These data were collated by date of injury and the patient's territorial local authority area (hereafter "district") of domicile. Sixty-seven districts currently exist in New Zealand. Four hundred and sixty-two hospitalisations with a district of "overseas/other" listed were included in the national descriptive statistics reported but excluded from analyses of the effects of temperature.

The date of injury for each incident was not necessarily the date of admission to hospital. Where multiple acts of self-harm contributed to a single hospitalisation, the date of the most recent act of self-harm was utilised as the date of injury. Furthermore, where the same act of self-harm appeared to have caused multiple hospitalisations (i.e., when the patient and the most recent date of injury was the same), only the first admission was counted. In less than 1% of the cases, the patient died in the hospital. These cases were included in our analyses.

The original source of the hospitalisation data was the National Minimum Dataset for hospital events. A problem with this dataset was inconsistency in terms of whether or not short emergency department stays were recorded, with different reporting practices used by different district health boards, as well as changes in reporting practices over time. We followed the practice of the Ministry of Health (2012) in excluding emergency department stays of less than two days from analysis, resulting in the exclusion of 24,906 incidents. With these incidents excluded, there were a total of 47,265 incidents of self-harm resulting in hospitalisation over the study period of 17 years.

4.3.3 Meteorological data.

Meteorological data was obtained from NIWA's virtual climate network rather than physical weather stations. This network covers New Zealand in an interpolated regular 5km grid (see NIWA, n.d.-a; Tait et al., 2006). The use of the virtual climate network avoided problems with missing data or with finding appropriate physical

---

[6] International Classification of Diseases, ninth edition, clinical modification (see Centre for Disease Control and Prevention, 2013).

stations to represent every district. The virtual weather station closest to the city centre of the largest town or urban area within each district was used to represent that district. A virtual climate station was not available for one district, the (offshore) Chatham Islands; the Chatham Islands automated weather station was used for this district. Mean daily temperatures were calculated by taking the mean of the daily minimum and daily maximum temperatures.

4.3.4 Population and demographic data.

Annual population data by district was obtained from Statistics New Zealand and interpolated to produce daily estimates. The percentages of the population in each district falling into various ethnicities and age groups were obtained from Statistics New Zealand (n.d.-a) for the censuses of 1996, 2001, and 2006. These percentages were then averaged across these time points within each district for use as controls in geographical comparison analyses.

4.3.5 Data analysis.

Data analysis was completed using R version 3.0.2 (R Core Team, 2013), with the package lme4 version 1.0–5 used for fitting linear mixed models (Bates et al., 2013). The Poisson model was used for most analyses. The Poisson model assumes that the conditional variance of the response variable is equal to the mean. Where variance in excess of the mean (i.e., overdispersion) was present the negative binomial model was used. Overdispersion was tested for by calculating the ratio of the Pearson chi-square fit statistic to the residual degrees of freedom, χ^2/df (Coxe, West, & Aiken, 2009). There was no evidence of temporal autocorrelation in any of the analyses used (e.g., lag 1 $r < 0.03$ in each case), meaning that an assumption that model errors were independent seemed reasonable. Confidence intervals for mixed models were calculated using the Wald method.

4.3.6 Ethical approval.

Ethical approval was obtained from the authors' host institution. Further information about the study's methods can be found in the Electronic Supplementary Materials.

4.4 Results

Excluding short emergency department stays, there were a total of 47,265 self-harm incidents resulting in hospitalisation in the study period. Of these, 64% were by females. The overall rate of self-harm incidents was 70.5 per 100,000 per annum.

4.4.1 Effects of geographical variation in temperature.

The mean temperature by district varied from 15.8°C in the Far North District to 9.5°C in Queenstown-Lakes. A simple correlation suggested that there was no evidence of a reliable relationship between self-harm rate per 100,000 and mean temperature across districts, r = .071, 95% CI [-.172, .306]. Negative binomial models were also estimated in which the number of self-harm incidents resulting in hospitalisation within each district over the entire study period was the response variable. In the first model, the only predictor variables were the mean temperature within each district over the entire study period and the logarithm[7] of mean population. In this model, the slope for geographical mean temperature was negative and very small.

The results differed somewhat when a second model was estimated with additional controls: radiation, the percentage of European, Asian and Māori residents in each district, and the percentages aged 15–39, 40–64 and 65 plus. Radiation was included as a control because the quantity of sunlight is a plausible third variable that might be related to both self-harm and temperature (see Doganay et al., 2003). In the controlled model, the point estimate of the effect of temperature was strongly positive, but the 95% confidence interval spanned zero. The coefficients for both models are displayed in Table 1.

---

[7] Given the use of a logarithmic link function as part of the negative binomial model, the log transformation allowed population size to have an additive effect. Specifically, one would expect the logarithm of population size to have a coefficient near 1, if self-harm incidence was directly proportional to population.

Table 1

Coefficients for Geographical Variation Models

	Uncontrolled model			Controlled model		
		95% CI			95% CI	
Coefficient	Est.	lower	upper	Est.	lower	upper
Intercept	6.654	6.557	6.754	6.581	6.472	6.694
Log population*	1.109	1.026	1.193	1.037	0.921	1.155
Temperature (°C)	-0.004	-0.062	0.054	0.050	-0.035	0.135
Radiation (MJ/m^2)				0.007	-0.098	0.112
Percentage European				-0.026	-0.060	0.005
Percentage Māori				-0.020	-0.045	0.004
Percentage Asian				-0.081	-0.166	0.004
Percentage aged 15–39				0.067	-0.002	0.137
Percentage aged 40–64				-0.001	-0.076	0.074
Percentage aged 65+				0.103	0.044	0.163

Notes. *Each variable was centred around its mean across districts, except for log population which was centred around the logarithm of the mean population estimate. Generalised linear model with negative binomial distribution and log link used. N = 67 districts.

4.4.2 Effect of seasonal variation in temperature.

There was relatively little seasonal variation in self-harm incidence, as is visible in Figure 8. The peak month was February (summer), with just 11.4% more self-harm incidents than the nadir in July (winter). For use in multivariable analyses, the seasonal norm temperature was operationalised as the mean temperature for each day of the 365 days of the calendar year, as averaged over the entire study period within each district. The total number of acts of self-harm resulting in hospitalisation occurring on the 17 occurrences of each day of the year for each district was then used as the response variable. February 29 was excluded, given that this day occurred just four times over the study period. A generalised linear mixed model with a Poisson distribution and log link was then fit, with population size controlled, and the intercept free to vary across districts. The resulting estimated effect of seasonal variation in temperature was modest (see Table 2), suggesting around 0.5% extra self-harm incidents resulting in hospitalisation per °C. Table 2 also includes estimates from a model with solar radiation controlled, resulting in the estimated effect of temperature becoming approximately zero. Seasonal variation in temperature and radiation were strongly correlated in our data (mean r = 0.8 across districts).

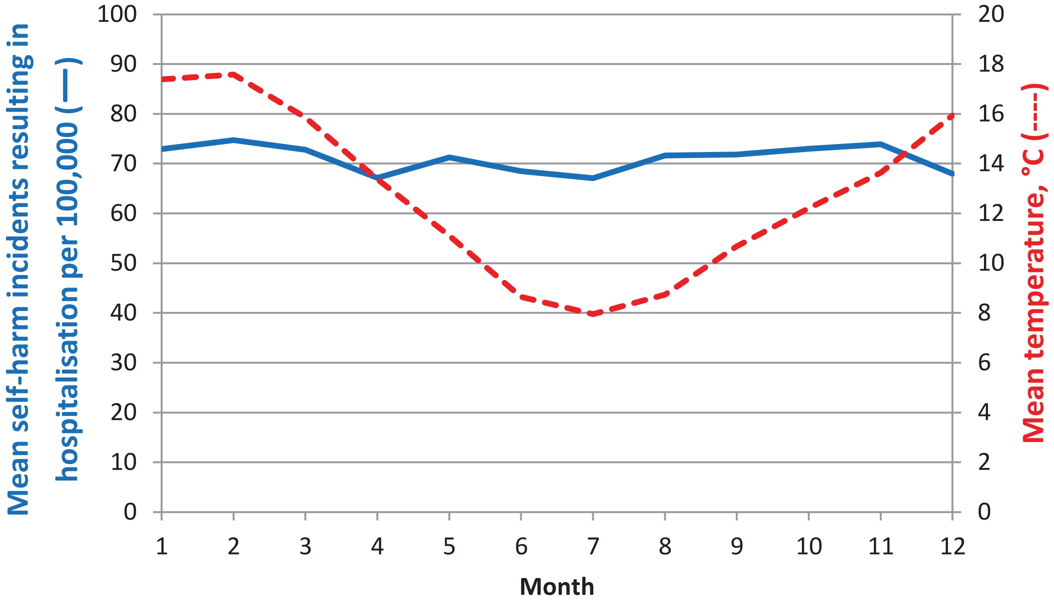

Figure 8. Self-harm incidents and temperature by month. The effect of number of days is controlled. Emergency department stays of one day or less are excluded.

Table 2

Coefficients for Seasonal Variation Models

	Model without radiation controlled			Model with radiation controlled		
		95% CI			95% CI	
Coefficient	Est.	Lower	Upper	Est.	Lower	Upper
Fixed effects						
Intercept	0.699	0.602	0.796	0.699	0.603	0.796
Seasonal norm temperature (°C)	0.005	0.002	0.008	0.001	-0.004	0.005
Log population* (log transformed)	1.106	1.028	1.184	1.106	1.028	1.184
Radiation (MJ/m^2)				0.003	1.9×10^{-4}	0.005
Random effects (SDs)						
Intercept | District	0.326	-	-	0.325	-	-

Notes. *Each variable was grand mean centred, apart from log population which was centred around the logarithm of the mean population estimate. Model: Poisson generalised linear mixed model with log link. N = 365 calendar days x 67 districts = 24,455. The response variable is the number of self-harm incidents occurring on a given day of the calendar year for a given district over the full 17 years of the study period.

4.4.3 Effects of irregular daily variation in temperature.

Irregular variation in temperature was captured by calculating temperature anomalies. These anomalies were the differences between the temperature observed on a given date in a particular district, and the seasonal norm temperature for that district and day of year (as defined previously). These anomalies were then entered into a mixed model, with the number of self-harm incidents resulting in hospitalisation occurring on each date and district as the response variable, and population controlled.

In the mixed model, a random intercept across both districts and years within districts was specified. This provided a flexible control for both regional differences in self-harm rate as well as for any potential time-varying confounds producing long-term trends in self-harm incidence. The coefficients for this model are shown in Table 3. Irregular daily variation in temperature appeared to have a moderate positive relationship with self-harm incidence, with the model implying an extra 0.7% acts of self-harm resulting in hospitalisation for every 1°C increase in temperature. Controlling for radiation in an alternative model resulted in no observable change to the coefficient for temperature (at 3 decimal places).

Table 3

Coefficients for Irregular Daily Variation Models

	Model without radiation controlled			Model with radiation controlled		
		95% CI			95% CI	
	Est.	Lower	Upper	Est.	Lower	Upper
Fixed effects						
Intercept	-2.197	-2.294	-2.101	-2.198	-2.294	-2.101
Temperature anomaly (°C)	0.007	0.003	0.011	0.007	0.003	0.012
Log Population*	1.059	0.982	1.135	1.058	0.981	1.135
Radiation (MJ/m^2)				0.002	4.9×10^{-4}	0.003
Random effects (SDs)						
Intercept | District	0.319	-	-	0.318	-	-
Intercept | District/Year	0.265	-	-	0.265	-	-

Notes. Each variable centred around its mean, except for log population which was centred around the logarithm of the mean population. Model: Poisson generalised linear mixed model with log link. Intercepts permitted to vary within districts, and within years within each district. *N* = 6209 calendar days x 67 districts less 69 days with missing temperature data = 415,934 for model without radiation controlled. *N* for model with radiation controlled = 415,857 due to an additional 77 cases with missing radiation measurements. The response variable is the number of self-harm incidents occurring on a given date in a specific region.

4.4.4 Trends in temperature and self-harm.

Figure 9 shows changes over time in the mean temperature and self-harm rate in New Zealand. Over the course of the study period, mean temperatures increased in New Zealand, with an average increase of 0.01°C per year. On the other hand, self-harm incidents resulting in hospitalisation declined markedly, with a linear model suggesting a downward trend of 1.8 fewer incidents of self-harm resulting in hospitalisation per 100,000 per year.

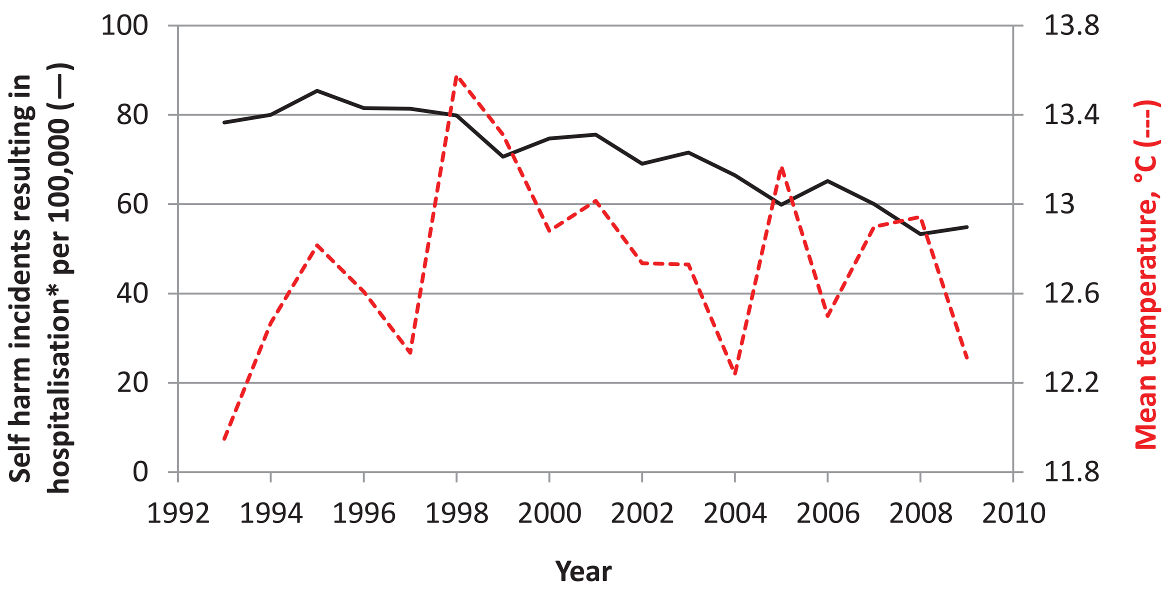

Figure 9. Trends in temperature and self-harm resulting in hospitalisation in New Zealand, 1993–2009. *Emergency department stays of one day or less are excluded.

4.5 Discussion

Our findings indicate that irregular daily variation in temperature had a small but positive relationship with the incidence of self-harm. In other words, days with hotter temperatures than the norm for that area and time of year tended to be associated with higher rates of self-harm. The size of this effect was about 0.7% extra incidents of self-harm resulting in hospitalisation per °C. This finding is similar to that of a recent study of suicides in New Zealand, which also found a small positive effect of irregular variation in temperature (Williams et al., 2015). Of the components examined, the effects of irregular daily temperature variations are the least likely to be confounded. This finding thus might suggest that higher daily temperatures occurring as part of climate change will exert a small increasing effect on self-harm incidence (around 0.4 extra acts of self-harm per 100,000 p.a. resulting in hospitalisation per °C, given the self-harm rate in 2009).

However, this conclusion is weakened somewhat by the lack of strong evidence for a positive effect of either seasonal or geographical variation in temperature. Seasonal variation in temperature did appear to have a positive relationship with self-harm incidence, but the relationship was small. This finding is consistent with a previous study that found limited (univariate) seasonality in suicide deaths in New Zealand (Yip et al., 1998) . Furthermore, the estimated effect of temperature was not robust to a control for radiation. Change in solar radiation is the major factor driving seasonality in temperatures, but seasonal variation in radiation and temperature are not perfectly correlated. This meant that the independent effects of the two variables could be estimated. The fact that radiation had a positive coefficient when both variables were included as predictors suggests that seasonal variation in radiation may exert some influence on self-harm incidence via a pathway other than by affecting temperature.

We were unable to confidently establish the direction or size of the relationship between geographical variation in temperature and the incidence of acts of self-harm resulting in hospitalisation. In a negative binomial model the estimated effect of geographical variation in temperature (i.e., mean temperature by district over the entire study period) had a very wide confidence interval, ranging from a strong negative effect to a strong positive one. This analysis was limited both by the small effective sample size in terms of geographical areas (just 67 districts) and the possibility of confounding by uncontrolled economic or demographic variables. This said, geographical comparison analyses do allow for at least some insight into the effects of long term differences in *climate*, as opposed to short-term variation in *weather*, and as such further research on this topic over larger geographical regions may be useful.

It was interesting to note that self-harm incidents resulting in hospitalisation (excluding short emergency department stays) declined over the study period, despite an increase in temperatures. It seems likely, however, that the downward trend in self-harm hospitalisations may be at least partially due to changes in medical practice and administration, such as increased use of community-based rather than inpatient mental health care (Ministry of Health, 2012).

This study did not attempt to uncover the causal mechanism of the relationship between temperature and self-harm. Little theoretical work has been attempted in this area, although the related topic of suicide seasonality has drawn far greater theoretical attention (for a review see Ajdacic-Gross et al., 2010). It is also worth acknowledging that our study focused only on the most serious cases of self-harm: those resulting in hospitalisation.

Overall, it remains challenging to predict how climate change will affect the incidence of self-harm in New Zealand. While the estimated effect of irregular daily variation in temperature appears to suggest that global (and regional) warming may increase self-harm incidence slightly, this conclusion would be more strongly justified were we able to show that seasonal and geographical variation had similar effects. A lack of consistency in the apparent effects of different components of variation in temperature raises the concern that humans may react differently to long-term, sustained differences in temperature than they do to short-term variation in temperature. This lack of consistency makes it more difficult to infer the future impact of a sustained increase in temperatures. None of the analyses in isolation are ideal as a basis for drawing inferences about climate change: The effects of irregular daily variation in temperature may not adequately approximate the effects of more sustained temperature changes, while analyses of the effects of seasonal and geographical variation are vulnerable to confounding. There is also the possibility that climate change may affect mental health via causal pathways not examined in this study (see Berry et al., 2010).

In conclusion, this study demonstrates that irregular daily variation in temperature is positively related to the incidence of self-harm resulting in hospitalisation in New Zealand. However, predicting how climate change will affect self-harm remains challenging. Future research focused on issues such as the relationship between geographical variation in temperature and self-harm, and the mechanism of the relationship between temperature and self-harm, may facilitate firmer conclusions about how climate change will affect self-harm rates.

4.6 Conclusions and Commentary

Study Three produced a rather familiar finding: As was the case for the studies of temperature effects on assault (Study One) and suicide (Study Two), there seemed to be a small positive effect of irregular daily variation in temperature on the incidence of acts of self-harm resulting in hospitalisation—albeit that this time the effect was about half the size as that of temperature on the incidence of suicides and assaults. Again, however, there was substantial uncertainty about both the size and direction of the effect of geographical variation in temperature, making it difficult to determine whether sustained exposure to warmer temperatures has a similar effect to that of random daily fluctuation in temperature. As in Study Two, the effect of *seasonal* variation in temperature—representing moderately sustained exposure to warmer or colder temperatures—was approximately zero. There was no sign of a negative effect of geographical variation in temperature, in contrast to the findings with respect to suicides in Study Two.

All three articles were able to produce only highly tentative conclusions about the effects of future climate change on violence. In each case, there was an apparent positive effect of the component of temperature variation least susceptible to confounding (irregular daily variation in temperature), but great uncertainty about whether more sustained exposure to warmer temperatures produced a similar effect. This meant in turn that only very tentative predictions about the future impact of climate change were possible in each case. Furthermore, the predictions that were attempted were predictions only about the *effect* of future increases in temperature, and not about the actual state of the world in the future. In other words, I did not at any stage attempt to predict the actual rate of violence at any particular time in the future.

The applied engagement with climate change research represented in these three empirical articles was valuable for the purposes of informing the methodological critique that follows. For example, actually engaging in research in this area demonstrated just how challenging it is to produce even tentative and conditional predictions about how human behaviour will be affected by a warming world. In the methodological critique, I will discuss how some of the features of psychologists' favoured methodological strategies make it particularly difficult for us to research climate change impacts in a way that is effective and useful. In some cases, the problematic research strategies are ones that I managed to avoid in my empirical studies, but this was certainly not always the case. The empirical studies were intended to inform and motivate the methodological critique that follows, and are not intended to be model examples of how psychologists should engage in climate change research.

5 Methodological Critique

This chapter forms the most important part of this thesis: an examination of the suitability of psychologists' mainstream methodological strategies for producing effective and useful research concerned with the impacts of climate change. In discussing methodological strategies I will draw especially on the testbed of research concerned with the relationship between temperature and acts of intra- and interpersonal violence. Both published literature and my own engagement in research in this area will be used to illustrate key methodological issues. This said, the aim is to examine the methodological strategies of psychologists in general—at least with respect to their suitability for studying the impacts of climate change—not just the strategies used by those who study the relationship between temperature and violence. At times I will therefore also discuss the use of methodological strategies in psychology more broadly.

As a preliminary, it is important to clarify which psychological impacts of climate change are of interest, and what types of causal pathways will be considered. Anthropogenic climate change is caused by increases in atmospheric greenhouse gas concentrations caused by human activities. The direct effect of this increase in greenhouse gas concentrations is an increase in global temperatures, but other meteorological and environmental variables are also likely to be affected, including sea levels, ice sheet cover, tropical storm intensity, precipitation patterns, and so forth (IPCC, 2013a). Both the mean values and frequency of extremes of environmental variables may be affected. These climatic changes will in turn affect ecological and human systems. Psychological impacts—effects on human behaviour, emotions or cognition—may be produced relatively directly: For example, natural disasters may cause an increased incidence of psychopathology (Rubonis & Bickman, 1991). Psychological impacts may also be produced more indirectly: For example, changes to climate may increase the incidence and prevalence of some diseases (McMichael, Woodruff, & Hales, 2006), subsequently impacting the mental health of some of those affected (Berry et al., 2010).

Correspondingly, some psychological impacts of climate change may be directly linked with fluctuations in meteorological variables that occur in the same period of time, while other effects may occur through pathways such that the link between meteorological event and ultimate impact is delayed and indirect. In this critique, I assume that in the short term psychological researchers will be particularly likely to focus on direct effects of climate change, such that the empirical analysis of relationships between meteorological and climatic events and behaviour will often be of interest. This is not to suggest that the study of the impacts of climate change can be reduced to the empirical analysis of the relationships between psychological and meteorological variables. Indeed, the study of other more complex causal impacts is also discussed at times in the arguments that follow.

5.1 The Impacts of Climate Change: Requirements for Effective and Useful Research

Before exploring the suitability of psychology's mainstream methodological strategies for producing effective and useful research concerned with the impacts of climate change, it is necessary to define what is actually required to make research in this area effective and useful. In this subsection I propose several requirements for research concerned with the psychological impacts of climate change to be effective and useful. The proposed list of requirements is certainly not exhaustive: Rather, the following requirements are ones that are particularly broad and crucial.

5.1.1 The need to identify existing effects of climate change.

The identification of psychological impacts of climate change that are already taking place is important for two major reasons. Firstly, in some cases existing impacts may be identified that require ameliorative action. For example, if research demonstrates that climate change is causing an increased frequency of mental illness in a specific region, practical action such as the provision of increased access to mental health services may be appropriate. Secondly, the identification of effects of climate change that are already occurring may inform predictions about the future consequences of climate change. For example, statistical models trained on observed data from the past may be used to generate predictions about the future. In this sense, research concerned with the existing effects of climate change is tied closely to research concerned with predicting the *future* impacts of climate change.

When identifying and studying the effects of climate change, it will also be important for psychologists to determine how effects *vary* across times, places, and populations. For example, as noted earlier, populations that are socially, politically and economically disadvantaged may be more vulnerable to climate change (IPCC, 2014b).

In other words, such disadvantaged populations may experience effects of climate change that differ in nature and magnitude from those experienced by less marginalised populations. It will be important for researchers to identify both the average effects of climate change on humans as well as how and why those effects may vary. Such information may help decision-makers to direct resources to ameliorate climate change impacts to the places where these resources are needed most.

5.1.2 The need to make predictions about changes to human behaviour in a warming world.

The Earth is already warming, but the increase in temperatures over the coming century is likely to be substantially greater than that observed thus far. This issue was discussed earlier in section 1.3.2. For psychological research concerned with the impacts of climate change to be useful, it will need to inform predictions about how human behaviour will change in the future as a result of climate change. Such predictions are important for several reasons.

Firstly, predictions about the impacts of climate change can inform attempts to mitigate climate change because they can increase understanding of the net future costs and benefits of anthropogenic greenhouse gas emissions. More negative impacts may imply the need for stronger and swifter action to reduce greenhouse gas emissions. Less negative impacts (or benefits of climate change) may imply less need for costly action.

Secondly, predictions about impacts can inform and motivate action to adapt to climate change impacts. For example, predictions of sea level rise due to global warming have led the government of the low-lying Maldives to set aside funds to assist in the eventual re-location of country’s population (Rameesh, 2008). Taking action to ameliorate or guard against some types of impacts may take a long period of time, so accurate predictions ahead of time can be of significant value.

In discussing the need for predictions it is possible to identify some characteristics that may make predictions more or less useful. Predictions may vary along several dimensions. Firstly, predictions may or may not be explicitly conditional on particular antecedents. For example, the IPCC provides temperature projections for a range of different greenhouse gas emissions scenarios (e.g., IPCC, 2013b). These predictions are provided without any explicit commentary as to which of the scenarios is most likely to actually transpire. The IPCC temperature projections are thus conditional on future anthropogenic greenhouse gas emissions. Secondly, predictions may be made with respect to the actual state of affairs in the world at some point in the future, or they may only refer to the *effect* that some anticipated event will have. Finally, predictions can vary in their degree of specificity. They may be specific about the magnitude, timing and location of some expected change or event, or they may simply indicate the direction of a likely change.

The type of prediction that is most valuable will depend somewhat on the researcher's goals. Predictions that are conditional on particular events (e.g., a prediction about behaviour that is conditional on the magnitude of global warming) may be useful for contributing toward an overall accounting of the costs and benefits of climate change. A prediction that refers only to the *effect* of an antecedent such as a temperature increase of a particular size—as opposed to an actual prediction of future behaviour—may similarly be useful when accounting for the overall costs and benefits of climate change. On the other hand, for a prediction to be useful for informing adaptation to climate change, it is more important that it actually predict the actual state of affairs in the world at a specific point in time (albeit with some unavoidable uncertainty).

While predictions that take both unconditional and conditional forms may be helpful in different settings, it seems likely that predictions about climate change impacts will almost always be more helpful when they are more specific about the magnitude, timing, and location of effects and events in the future. As such, the production of specific and quantitative predictions—as opposed to statements suggesting only the general direction of effects or changes—is an important activity when performing research concerned with the psychological impacts of climate change. Such specific predictions will be particularly useful if they include information about how predicted effects may vary across populations, places, and times.

5.1.3 The need to accurately estimate the uncertainty surrounding conclusions.

The management and communication of uncertainty is an activity of major importance throughout virtually all areas of science. Friedman, Dunwoody, and Rogers (1999) describe these activities as a "hallmark of good science" (p. vii). Modelling and communicating uncertainty is especially important when drawing conclusions about climate change, because the resulting conclusions may be used to make real-life decisions that have important and widespread consequences. These decisions need to be informed by an adequate understanding of the quantity of uncertainty about the likely consequences of different courses of action.

Estimation of the uncertainty surrounding statements about the impacts of climate change is particularly important because of the connection between uncertainty and *risk* (see Bammer & Smithson, 2008). The problem of low-probability high-cost events is particularly germane to the context of climate change; such events can have a very large effect on the estimated net costs of climate change (see Weitzman, 2009). The appropriate practical response to an effect whose size can be predicted with minimal uncertainty may be very different from the appropriate response to an effect whose size is subject to a great deal of uncertainty. In some cases, the presence of substantial uncertainty may imply the need to be prepared for less likely but more severe events. This means that it is crucial that psychologists are able to estimate and report as accurately as possible the quantity of uncertainty surrounding their predictions about the impacts of climate change.

5.1.4 The need to clearly communicate findings to policymakers and other non-psychologists.

Relatedly, a requirement for research on the impacts of climate change to be effective and useful is that the findings of the research are communicated to those who can use the information to help humans mitigate or adapt to climate change. As discussed earlier (see section 1.3.4), psychologists engaging in research concerned with climate change may need to effectively communicate their findings to non-psychologists, including policymakers, physical scientists, and laypeople. The study of communications about climate change is indeed becoming an important research area in its own right (Moser, 2010). Effective communication about climate change may mean communicating information directly to policymakers, or via venues likely to receive attention from policymakers and the general public, such as the mainstream media. It may also mean ensuring that findings are disseminated in the IPCC reports, which are one of the primary conduits for research findings about climate change to be communicated to policymakers. Indeed, each section of each IPCC report is explicitly summarised into a summary for policymakers chapter (e.g., IPCC, 2014a). In order to ensure that psychological studies concerned with climate change are included in IPCC reports, the authors of such studies may need to take action to bring this research to the attention of the IPCC authors, who are in most cases physical rather than social scientists. One useful strategy may to publish psychological research in multidisciplinary journals concerned with climate change (e.g., *Climatic Change*, the venue for two of the articles published based on this thesis).

Bringing research to the attention of those who can use it is not enough, however. It is crucial that researchers are able to clearly and accurately communicate their conclusions, and the uncertainty relating to those conclusions, in a manner that the intended audience can understand. For example, laypeople may interpret the language used to communicate uncertainty in scientific documents in a manner that differs significantly from the interpretation intended by the authors. This can be a problem even when the terms used to define uncertainty are clearly and directly defined, as in the IPCC reports (see Budescu, Broomell, & Por, 2009). It is likely to be an even greater problem when researchers use indirect and counterintuitive methods to communicate uncertainty, such as *p* values.

5.2 The Suitability of Psychologists' Methodological Strategies for Meeting the Above Requirements

Psychologists are starting to engage in research concerned with the impacts of climate change (e.g., Anderson & DeLisi, 2011; Doherty & Clayton, 2011; Page & Howard, 2010; Reser, Morrissey, & Ellul, 2011; Swim et al., 2011). Unfortunately, there are several key problems with the methodological strategies generally utilised in mainstream quantitative psychology that may make it difficult to produce effective and useful research concerned with the impacts of climate change. These problems are discussed in the critique that follows.

5.2.1 Use of measurement and data analysis strategies that are unconducive to clearly communicating the sizes of effects.

When communicating conclusions about the impacts of climate change, it is important that researchers are able to convey information about the size of effects. A statement that an effect is or will be in a particular direction (positive or negative) is of little use without clear information about the magnitude of the effect. In fact, the magnitude of effects is just as important as their direction. The appropriate policy response to a prediction that a particular effect will be very small might be the same regardless of whether the effect is positive or negative (because no substantive response is required). On the other hand, a conclusion that an effect will be large and negative might require a drastically different response than if only a small negative effect was expected. More generally, a conclusion that refers solely to the direction of an effect produces very limited information about the size of the effect. For example, a statement that climate change will *increase* the incidence of some mental disorders is equivalent only to saying that the effect of climate change on the incidence of the disorder will be of a size somewhere between zero and an infinitely large positive effect. As such, it is important to provide estimates of the specific size of effects, and not just their direction. Such estimates will be valuable even if they are subject to substantial uncertainty (e.g., even if they take the form of very wide interval estimates of effect size). Unfortunately, there are two major reasons why psychologists often struggle to clearly and accurately convey information about the sizes of effects.

5.2.1.1 Measurement strategies that damage the meaningfulness of effect size estimates.

The first reason is that psychologists often use measurement tools that do not have clearly defined and meaningful units of measurement. For example, consider a study of participants affected by a natural disaster (a flood), a man-made disaster (a toxic landfill) and a comparison group (Baum, Fleming, Israel, & O'Keeffe, 1992). This is a study of particular interest when considering how disasters caused by man-made changes to the climate might in turn affect mental health. Baum et al. found that flood-affected participants had a mean score of 21 on the Symptom Checklist-90-Revised (SCL-90). On the other hand, those affected by the man-made disaster had a mean score of 38, and comparison participants had a mean score of 21. This finding suggested that the man-made disaster had a more severe effect on mental health than the natural disaster (the flood). The problem here is conveying what the average difference of 17 points between comparison participants and those affected by a man-made disaster actually *means*. Understanding the practical importance of this difference requires not only substantive knowledge about the topic at hand, but also a great deal of knowledge about incidental features of the measurement process. These features include the content of items and response options, the number of items, the number of possible responses to each item, the scoring scheme used to translate item responses into numeric values, and the method used to combine responses into an overall summary score. Many of these features are likely to differ across studies and measurement devices.

Similar problems have occurred in the temperature-violence testbed: For example, Anderson et al. (1995, Experiment 2) reported the effect of temperature on state hostility in a laboratory experiment. State hostility was measured via a 35-item Likert scale with a possible score range of 35–175. The resulting slope of a 0.8-point increase in state hostility for every 1°C was again not a very effective way to communicate effect size. Interpreting this figure requires specific knowledge not just about the construct being measured, but also about a number of incidental features of the measurement process such as those discussed in the paragraph above.

Furthermore, the meaning of a 0.8-point difference on a psychometric scale such as that used by Anderson and colleagues (1995) may even vary at different points on each measurement scale. There is no evidence that psychological attributes are themselves actually quantitative[8] (see Michell, 1997, 1999, 2000). In other words, there is no evidence that Hölder's axioms of quantity (see Michell & Ernst, 1996) actually apply to any psychological attribute. Conceptually, this means that the differences between various levels of a psychological attribute itself may not just be homogeneous differences on a quantitative scale, but rather may qualitatively differ from one another. For example, correctly answering a moderately difficult question on an intelligence test may require the possession of a specific skill or item of knowledge that is not required to answer an easier question. Answering an even more difficult question may require the additional possession of yet another (qualitatively different) skill or item of knowledge. The difference between levels of a psychological attribute such as intelligence may well therefore represent not just differences in quantity, but differences in kind. In turn, this means that scales intended to measure psychological attributes will usually produce scores that have (at best) an ordinal level of measurement. Ordinal scales lack a unit of measurement that has a constant meaning across different levels of the measured attribute[9]. In contrast, in climatology, a difference of 1°C in temperature has the same

---

[8] Michell (1997, 1999, 2000) also argues that only quantitative variables are measurable, and that therefore psychologists cannot claim to be able to measure psychological attributes. For the sake of convenience, in this thesis I use the term *measurement* more broadly, in line with Stevens's (1946) very loose definition of measurement as "the assignment of numerals to objects or events according to rules" (p. 677).

[9] An ordinal scale arises when we have observed the *ordering* of objects with respect to their values of some attribute (e.g., A is greater than B), but when we have not been able to observe or compare the sizes of the *differences* between objects (see Michell, 1986; Stevens, 1946).

meaning[10] regardless of the point on the temperature scale at which this difference occurs and regardless of the measurement procedure used to identify the difference.

The fact that psychological test scores do not have established, clearly defined units of measurement that apply across measurement devices and across different levels of a measured attribute means that it is especially difficult for psychologists to convey information about effect size. Specifically, it is difficult to communicate the size of the effect of a particular change in conditions on a behavioural dependent variable if that dependent variable has no clearly defined unit of measurement. For example, a simple unstandardised effect size such as a mean difference across two conditions conveys clear information about the size of the effect only if this mean difference is expressed on a scale with clear units of measurement. If a behavioural variable itself is not measured on a scale with a clear and interpretable unit of measurement, then the unstandardised[11] size of a particular effect on that variable will likewise have no clear and interpretable unit of measurement.

The achievement of true quantitative measurement of psychological attributes with clearly defined units of measurement, is obviously a difficult task—and potentially an outright impossible one, if it indeed transpires that psychological attributes are not themselves quantitative. Interestingly, however, psychologists doing work relevant to climate change have sometimes chosen to use measurement scales without clearly defined units of measurement even for measuring *meteorological* variables. This is despite the fact that the use of meteorological variables mean that quantitative measurements with clearly defined units are available. For example, Anderson and Anderson (1996, Study 2) investigated the relationship between geographical variation in temperature and violence. In their study, they operationalised temperature by transforming and summing four indicators of the climate of US cities in 1980. These indicators were the number of hot days (≥ 32.2°C), the number of heating degree days[12],

[10] An example of how the meaning of a 1°C temperature difference remains the same across the temperature scale can be seen in the fact that the energy required to heat a given object by 1°C remains the same regardless of the initial temperature of the object (see Senese, 2010). A more detailed discussion of the evidence that temperature is a quantitative variable—and thus can be measured on a scale with a unit of measurement that has a constant meaning—can be found in Sherry (2011).

[11] The alternative strategy of using standardised effect sizes is critiqued in the next subsection.

[12] The number of heating or cooling degree days for a particular day is defined as the difference between that day's temperature and a reference "comfortable" temperature, generally 18.3°C. For example, a day with an observed mean temperature of 24.3°C would have a measurement of 6 cooling

the number of cold days (≤ 0°C), and the number of cooling degree days. Each indicator was converted to *z* scores, the latter two indicators were reverse coded, and the *z* scores were summed to produce a "hotscale" variable. Despite the fact that temperature has clearly defined units of measurement, this hotscale variable does not, meaning that the study was unable to report an estimate of the effect of a unit increase in temperature on violent crime.

In sum, problematic measurement strategies used in psychology can prevent the clear reporting of effect size. The most widespread and problematic strategy is the use of psychological test scores without clearly defined units of measurement (e.g., Anderson et al., 1995; Baum et al., 1992), but problems can occur even when quantitative scales with clear units of measurement are available (as in Anderson & Anderson, 1996).

5.2.1.2 Data analysis strategies that hinder the communication of effect sizes.

Another reason why psychologists may struggle to effectively convey information about the magnitude of climate change impacts is that psychologists have a predilection for using statistical analyses that hinder the clear reporting of effect size. The most obvious case of such a problem is when results are reported purely in terms of statistical significance (i.e., *p* values), without further information about effect size. *p* values, of course, do not directly communicate effect size[13] (see Chow, 1988; Jacobson & Truax, 1991; Kirk, 1996; Nakagawa & Cuthill, 2007). Within the testbed of temperature-violence literature, some older studies relied primarily on statistical significance tests for conveying information about effects (Baron & Bell, 1975; Bell & Baron, 1976). However, it fortunately seems to be a relatively rare scenario in this research area that a researcher reports a finding purely in terms of statistical significance without any standardised or unstandardised measure of effect size.

A more common problem, however, is the reporting of relationships purely or primarily in terms of standardised effect sizes such as correlations (Barker et al., 1994; Doganay et al., 2003; Lester, 1999) or standardised regression coefficients (e.g., Preti, 1998). Within the field of research concerned with temperature and violence, geographical comparison studies in specific have often tended to use standardised effect

degree days and 0 heating degree days. Heating and cooling degree days can then be summed across longer periods of time, such as a year.

[13] The *p* value for a particular relationship is of course *partially* determined by the size of the effect observed. It is nevertheless a poor way to communicate effect size.

sizes. For example, Anderson and Anderson's (1996) comparison of violent crimes across cities in the USA reported correlations and partial correlations between crime and climatic variables. A subsequent re-analysis of Anderson and Anderson's data used structural equation modelling with standardised coefficients (Anderson et al., 2000). Similarly, two international studies of the relationship between suicide rates, homicide rates and temperature relied purely on correlation coefficients (Robbins et al., 1972; Rotton, 1986). Standardised effect sizes convey relationships using the standard deviations of all or some of the variables in question as the units of measurement. For example, the Pearson's correlation coefficient r indicates the expected standard deviation increase in one variable given a standard deviation increase in another variable. Other well-known standardised effect sizes include the standardised regression coefficient Beta, Cohen's d, and eta-squared.

Standardised effect sizes are attractive to psychologists because they appear to help to convey the magnitude of relationships between variables that lack clear units of measurement. Instead of having to interpret an effect size expressed in measurement units that do not themselves have a clear meaning, researchers are able to rely on effect sizes expressed in a standardised metric, with rule-of-thumb guidelines available for defining "large", "medium", and "small" effects (e.g., see J. Cohen, 1988, pp. 79–80). Unfortunately, however, standardised effect sizes do not necessarily provide clear information about the practical significance of effects. Baguley (2010) gives the example of the .24 correlation between earnings and height in the USA (based on Gelman & Hill, 2007). This correlation would be considered to somewhere between "small" (r = .1) and "medium" (r = .3) according to the popular guidelines of J. Cohen (1988, pp. 79–80). Yet the corresponding unstandardised regression coefficient of $\hat{\beta}$ = 1256 suggests that this interpretation might be problematic: In fact, every extra inch in height was associated with a USD1256 increase in annual earnings. This effect is surprisingly large and of potential practical importance, despite the rather unimpressive correlation coefficient. In general, attempting to convey effect size without any reference to the original scale of the variables analysed may result in misleading communications.

A more specific problem with standardised effect sizes is that because they are scaled with reference to the variability of the variables analysed, they are dependent not only on the sizes of effects but also on methodological strategies that impact variability. Such strategies may be as arbitrary as the choice to use a particular level of aggregation.

For example, consider two different methods for calculating the relationship between seasonal variation and assault incidence within a geographical region. One method is to calculate the average temperature for each of the 365 days of the calendar year (across the entire period period) and correlate these temperatures with the assault rate for each of the 365 *days* of the calendar year, again as averaged across the study period. Based on the recorded assaults data examined in Study One[14], this leads to a correlation of *r*(363) = .40 in the region of Auckland, for example. But a second way to approach the same question is to apply the same strategy but aggregate temperature and assault rates over the 12 *months* of the year. This simple change in aggregation level—without any change to the actual data analysed—produces a substantially larger correlation of .62 for the same dataset. This is despite the fact that the estimated (unstandardised) effect of temperature is virtually the same in both cases, at around 1.5% extra assaults for every 1°C increase. The reason for the larger correlation when monthly aggregation is used is that aggregation over months removes much of the variability in assaults that is *not* due to seasonal variation in temperature. Indeed, some of the studies reporting the largest correlations between temperature and acts of violence have been based on data aggregated by month in this fashion (e.g., Doganay et al., 2003; Preti, 1997).

Fairly arbitrary decisions about research methods can thus have an important effect on standardised effect estimates. From a statistical perspective there is an even more concerning problem, however: that of bias in standardised effect size estimates. Recall that standardised effect sizes typically report the sizes of estimated effects scaled with respect to the variability of the dependent variable (or of both the independent variable and the dependent variable). Perhaps the simplest example is Cohen's *d*, which is calculated as:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

Where $\bar{x}_1$ is the mean of the dependent variable in group 1, $\bar{x}_2$ is the mean of the dependent variable in group 2, and *s* is the pooled standard deviation of the dependent variable (see J. Cohen, 1988). The value of *d* calculated from sample data can be interpreted as an estimate of the population parameter δ: that is, the difference between the two *population* means divided by the standard deviation of the dependent variable, as pooled across the two populations. The problem is that if the pooled sample standard

[14] These results are not reported in Study One, and were calculated for the purpose of this specific illustration.

deviation s is a biased estimate of the population standard deviation σ, then all else being equal, d will obviously be a biased[15] estimate of δ.

Crucially, the pooled sample standard deviation will be a downwardly biased estimate of the population standard deviation σ in realistic conditions. Researchers (especially those conducting experiments) often *deliberately* seek to remove the effects of extraneous variables on the dependent variable, thus reducing its standard deviation. Indeed, methodological texts often explicitly recommend seeking to remove the effects of variables other than the independent variable on the dependent variable (e.g., Cozby & Bates, 2015; Price, 2012). This may be accomplished by conducting experiments in tightly controlled settings and/or by selecting samples that are relatively homogeneous. These strategies are reasonably well justified, since they improve statistical power, reduce the threat of confounding, and do *not* lead to bias of unstandardised effect sizes. For example, for the sample mean difference $\bar{x}_1 - \bar{x}_2$ to be an unbiased estimate of the population mean difference $\mu_1 - \mu_2$, we require only one assumption: that the error terms (the distributions of the differences between the sample mean and the observations within each population) have an expected value of zero. Yet even when the unstandardised effect size is unbiased, design strategies that remove sources of variation in the response variable other than the predictor variable will still result in a Cohen's d effect size that is an upwardly biased estimate of the population effect size δ. This is trivially the case, given that Cohen's d is calculated by dividing the mean difference by the pooled standard deviation. If the standard deviation is biased downward, the d estimate will be biased upward. The magnitude of this bias could be serious: If the pooled sample standard deviation has a 20% downward bias, for example, the resulting upward bias of Cohen's d is approximately $\frac{1}{1-0.2} - 1$ = 25%. Similar problems apply to other standardised effect sizes that are scaled with respect to the standard deviations of the variables involved. If a standardised effect size is a biased estimate, then its value for communicating results is compromised.

[15] Admittedly, standardised effect sizes including Cohen's d, Pearson correlations and others are actually biased estimates even when the sample standard deviation s is an unbiased estimate of σ (Fisher, 1921; Rosenthal, 1994). This well-known bias is greater with smaller samples, and occurs due to the asymmetric distribution of these statistics, but is small in magnitude and readily corrected, as Rosenthal notes. On the other hand, bias in the estimate of d caused by a biased standard deviation estimate (as discussed in the text above) may be much more severe.

A pragmatist response to this critique might be that even if a standardised effect size is a biased estimate of the corresponding population parameter, it is still a useful descriptive statistic representing the size of an effect in a particular sample. A problem with this point of view is that if a standardised effect size depends heavily on arbitrary methodological decisions made by the researcher that impact variability, it is unclear why that standardised effect size would represent any more of a meaningful or useful estimate of the effect size in the sample than an unstandardised estimate. Furthermore, a large part of the motivation for reporting standardised effect sizes is the aim of reporting results in a metric that can be compared and averaged across studies (as in a meta-analysis). It is surely only appropriate to average the effect sizes reported by a number of studies if all of these effect sizes represent estimates of the same underlying quantity. If, in fact, a selection of effect sizes from different studies represent estimates with quite different underlying metrics—because the unit of measurement for the effect size reported in each individual study is at least partially determined by idiosyncratic features of that study that impact variability—then these effect sizes are not multiple estimates of a single quantity. Rather, they are all estimates of different quantities with different units of measurement, and averaging them is little more sensible than averaging unstandardised estimates of effect size across studies using scales with completely different units of measurement. If standardised effect sizes do not facilitate informative meta-analyses, then one of the primary rationales for their reporting is weakened. An alternative strategy for performing meta-analysis is discussed in section 6.2.

Another problem with standardised effect sizes that is particularly problematic in the research area concerned with the effects of temperature on violence is that it is difficult (without additional information) to convert a correlation between measures of temperature and violence into a prediction about the effect of an increase in temperature of a particular magnitude. This makes such standardised effect sizes rather unsuitable for use when making predictions about climate change. Of course, standardised effect sizes can be back-transformed into unstandardised effect sizes (e.g., regression coefficients) if the standard deviations of the variables are reported, and thus used to generate predictions. However, the standard deviations of the unstandardised variables were not reported in several of the studies using standardised effect sizes that were mentioned above (e.g., Anderson & Anderson, 1996; Anderson et al., 2000; Robbins et al., 1972; Rotton, 1986). Furthermore, any requirement on the reader to perform further calculations to make research outputs useful is highly undesirable, especially in an applied area where communication of information to policymakers and other non-scientists is an important goal.

In sum, both problematic measurement strategies as well as problematic statistical strategies can hamper the clear reporting of effect size. Measurement strategies that hinder the clear reporting of effect size include the use of variables without clearly defined units of measurement, while data analysis strategies that cause the same problem include the reporting of standardised rather than unstandardised effect sizes. The use of standardised effect sizes is a particularly problematic strategy for a variety of reasons, including the possibility of communicating misleading information about the practical significance of an effect, the fact that the size of the effect reported will depend heavily on arbitrary decisions made by the researcher, the upward bias of standardised effect sizes in realistic conditions, and the fact that standardised effect sizes are of limited use as a basis for predictions.

5.2.2 Use of theory and analysis strategies that limit our capacity to predict future events.

As indicated earlier in this chapter, psychologists studying climate change impacts will need to make predictions, where the predicted events may take place decades or even centuries in the future. Even given the production and communication of clear information about the size of climate change impacts, problems with psychologists' favoured methodological strategies may still result in difficulties in producing accurate and useful predictions. These problems include a general neglect of the timescale of effects and a lack of development of theory that can support the making of quantitative predictions.

5.2.2.1 Neglect of the timescale of effects.

While climatic warming may accelerate this century, anthropogenic climate change remains a problem that will evolve over a long period of time, with the possibility of changes that may not be reversible for centuries or millennia (see Collins et al., 2013). Furthermore, conclusions about the effects of a long-term, sustained change in climate may often need to be drawn based primarily on data showing how humans react to short-term variation in climatic conditions. It is therefore crucial that psychologists are able to account for the role of time when making conclusions about the psychological impacts of climate change. Psychologists will need to identify the timeframe of specific causal processes and how effects differ depending on the timeframe of exposure to particular causal antecedents. Of particular concern is the possibility that long-term variation in climatic factors could produce effects that are quite distinct from those of short-term variation. This may occur, for example, as a result of adaptation to sustained climate change. Unfortunately, the role of time is rarely accounted for in either psychological theory or data analysis. Certainly this is the case within the temperature-violence testbed.

5.2.2.1.1 Time: Ignored in theory.

Neglect of the role of time is an unfortunate feature of the majority of the prominent theories in the temperature-violence testbed. For example, the general affective aggression model (Anderson et al., 1995; Lindsay & Anderson, 2000) suggests that situational variables such as uncomfortable temperatures can elicit hostile cognitions, physiological arousal, and angry emotions, thus increasing the probability of aggressive behaviour. However, the theory does not specify how long this process takes or what difference the time of exposure to uncomfortable temperatures makes. Similarly, the negative affect escape model (Bell & Baron, 1976) suggests that moderate levels of negative affect, as may be evoked by uncomfortable temperatures, can increase aggressive behaviour, and that very high levels of negative affect prompt escape from a stressful situation. But again, the theory says nothing about what effect the time of exposure to negative affect has on the probability of aggressive behaviour. Along with the fact that these theories do not make any quantitative statements about the size of the effects involved, the lack of any description of the role of time makes it difficult to use these theories to make predictions about climate change. In a more general sense, theories of behaviour that specify causal relationships between variables, but do not describe how and when causal processes take place in real time, may only ever provide a fairly limited description of the mechanisms underlying human behaviour.

Theorising in some areas of psychology does have an explicit or implicit focus on time. Developmental psychology is an obvious example of such a research area; another is that of dynamical approaches to cognition (van Gelder & Port, 1995). Nevertheless, many major areas of psychology have given little attention to the role of time. Roe (2008) suggests that the role of time is “barely acknowledged in the psychological literature" (p. 37). Time has a much more important role in the theory of physical sciences such as climatology. For example, time is one of the independent variables in the Navier-Stokes equations (NASA, n.d.), these equations being the basis of general circulation models used to simulate the past and future climate (Fraedrich, Aigner, Kirk, & Lunkeit, 2006). If psychological theory is to be useful for drawing conclusions about the impacts of climate change, it will similarly need to be developed in such a way as to incorporate the important role of time in causal processes.

5.2.2.1.2 Time: Ignored in analysis.

Another research activity in which time is often ignored by psychologists is in the performing of data analysis. In general, analyses of the effects of temperature on intra- and interpersonal violence have been conducted with little focus on how long the relevant causal processes take. Instead, it is generally assumed that cause (temperature variation) and effect (violence) occur essentially simultaneously, regardless of the duration of the time periods analysed. Occasional exceptions to this rule primarily take the form of exploratory analyses of lagged temperature effects (e.g., Y. Kim et al., 2011; Likhvar et al., 2011, and Studies One and Two in this thesis). These analyses may be useful for providing some indication of the timeframe of causal effects, but are primarily useful only for identifying slightly-delayed causal effects: for example, an effect of temperature variation that is delayed by a lag of a day or two. Such analyses are less useful for identifying causal processes that occur over longer timeframes.

Another potentially valuable strategy for studying how effects differ according to the time of exposure to climatic phenomena is the study of different forms of variation in climate and weather. These forms of variation include irregular/random daily variation (short-term exposure), seasonal variation (medium-term exposure), trends (longer-term exposure), or geographical variation (essentially permanent climatic differences between areas). The similarities and differences between the effects of temperature variation on intra- and interpersonal violence across different timescales were given a great deal of attention in the empirical studies reported in this thesis. The study of the effects of variation in climatic factors across different timeframes is important because it can begin to answer questions about how longer-term exposure to an altered climate affects human behaviour. For example, as noted in Study Two, short-term irregular variation in temperature seems to produce increased suicide risk, while geographical variation in temperature (that is, long-term exposure to higher temperatures) seems to be associated with *reduced* suicide risk. This finding very tentatively suggests that longer-term exposure to higher temperatures may produce adaptive responses that reduce suicide risk.

The idea that the effects of different components of variation in temperature might have different effects has not been recognised very clearly in the literature concerned with temperature effects on violence. Literature reviews by the Anderson group (Anderson, 1989; Anderson et al., 2000; Anderson & DeLisi, 2011) have discussed the consistency of the direction of the effect of temperature across geographical, time period (temporal), and experimental studies, noting that these studies have used different research *methods.* However, their series of reviews did not discuss the idea that geographical, temporal, and experimental studies might be informative as to the effects of different durations of exposure to climatic phenomena. In the area of literature concerned with temperature effects on intrapersonal violence there has been even less recognition of the possible differences between the effects of different forms of variation in temperature. For example, Deisenhammer's (2003) review of studies of the relationship between weather and suicide made no distinction between geographical and temporal studies, thus resulting in a conclusion that temperature's effect on suicide has been inconsistent across studies. In reality, the effect found has been relatively consistent once the type of variation studied is taken into account (as discussed in Study Two, section 3.2).

5.2.2.2 Lack of theory development that facilitates prediction.

In climatology, theory is drawn on heavily when producing predictions. The models used by climatologists to predict future climatic changes are both *mathematical* and *theoretical*. In other words, they take the form of mathematical equations that can be used to make predictions (as do statistical models), but the parameters in these equations are specified by theory (unlike the parameters in statistical models, which are estimated from empirical data). For example, general circulation models—an important type of model used for predicting the global climate—make predictions based on the Navier-Stokes equations (see Fraedrich et al., 2006). The Navier-Stokes equations arise from applying Newton's second law to fluid motion.

Psychologists unfortunately have developed little in the way of theory that takes a mathematical form and is thus capable of producing quantitative predictions (but see Luce, 1999). Rather, psychological theory typically takes a verbal form (Myung & Pitt, 2002). Specifically, psychological theory is often couched in the form of a series of statements about relationships between variables. These statements may indicate the direction of the relationships between variables, and so too the presence of any hypothesised mediating or moderating relationships, but the size of the relationships (the model parameters) are *not* specified. Naturally, theories that take this non-quantitative form cannot be used in isolation to produce quantitative predictions. To make quantitative predictions, psychologists generally have to use statistical models, with parameters estimated by fitting models to empirical data. Certainly, psychologists attempting to predict climate change impacts can still use verbal theory (and abductive inference more generally[16]) to inform the specification of statistical models that are capable of making predictions. However, without the development of mathematical theory, the generation of predictions using such statistical models will always depend on the availability of suitable empirical data with which to train (i.e., estimate) the statistical models. This requirement of the presence of suitable empirical data is a problem for psychologists studying climate change impacts, who will generally have access to very limited data showing how human behaviour has responded to sustained climate change in the past.

Even setting aside the need for empirical data with which to train statistical models that are based on theory, theories will vary in their degree to which they can usefully inform the specification of statistical models used to make predictions. I turn to this problem next.

5.2.2.2.1 Specification of causal effects that cannot be estimated.

Importantly, not all psychological theories—even the ones that do imply some testable predictions—will be capable of guiding the development of statistical models that can be applied to generate useful predictions. For example, a theory that is constituted of statements about causal effects between variables[17] will need to specify causal effects that can be empirically estimated in the specific context and in relation to the specific behaviours for which predictions are sought.

An example may help to clarify this point. One of the causal effects specified in the general affective aggression model (Anderson et al., 1995; Lindsay & Anderson,

[16] Abductive inference, roughly speaking, is the process of generating hypotheses that explain past observations (see Haig, 2005).

[17] This is not the only type of psychological theory, but it is the most popular type; see Jaccard and Jacoby (2010).

2000) is an effect of hostile thoughts on the likelihood of a choice to aggress. It is possible to estimate this causal effect for at least some operational definitions of the specified constructs. For example, in a laboratory setting, one could measure or even manipulate hostile thoughts and measure aggression using a proxy measure such as the intensity and duration of noise blasts delivered to an opponent in a competitive game (as in Lindsay & Anderson, 2000). It is possible then to estimate the size of the relationship between these two variables. But information about the size of the effect of hostile thoughts on the intensity and duration of noise blasts administered by participants does not tell us about the size of the effect of hostile cognitions on the incidence of an aggressive behaviour of real substantive interest (e.g., assaults). Indeed, it would be much more difficult to empirically estimate the effect of temporal variation in hostile cognition levels on actual assault rates. As such, even if the general affective aggression model has identified a causal effect that exists and that can be estimated in laboratory settings, it has not necessarily identified an effect that can actually be used to generate useful quantitative predictions.

Similarly, several other theories that purportedly explain the effect of temperature on violence hypothesise causal effects that can realistically only be estimated in laboratory environments with artificial proxies for the behaviours of real interest. Like the general affective aggression model, the negative affect escape model (Bell & Baron, 1976) invokes emotional variables (such as negative affect) whose effects on aggressive behaviour can be measured by using proxies for aggression in laboratory contexts. However, the effects of negative affect on more important acts of aggression such as assault would be very difficult to estimate for practical and ethical reasons. Routine activity theory (L. E. Cohen & Felson, 1979) specifies a causal effect of opportunity on crime as its central claim, but the overall level of opportunity for crime within a geographical area and time period is virtually impossible to define, let alone measure. In sum, while the development of theory may have an important role in helping to produce predictions, much psychological theory—and most of the theory in the temperature-violence literature—may not have the necessary characteristics to support the making of practical and useful predictions.

5.2.3 Inadequate modelling and reporting of uncertainty.

A final broad feature of psychologists' conventional approach to research that is likely to cause problems when studying climate change impacts is an insufficient focus on modelling and reporting information about uncertainty. The problems with how psychologists address uncertainty can be illustrated by comparison with how uncertainty is addressed in climatology. In climatology, the estimation and reporting of uncertainty arising from many different sources is given a great deal of attention. Climatologists attempting to make predictions about the future estimate many simulation models, with the models incorporating different assumptions about initial conditions, different assumptions about future anthropogenic emissions, and having different results due to random internal variability. By studying the variability of the estimates produced, and not just their central tendency, climatologists can both provide projections of the future climate along with information about the uncertainty surrounding these projections. IPCC climate projections are typically provided in the form of interval estimates achieved by incorporating the predictions of many models, thus addressing uncertainty occurring due to a range of sources (see for example, IPCC, 2013a).

The reporting of uncertainty in psychology is conducted differently, and is generally less comprehensive than in climatology. Specifically, the quantitative tools for reporting uncertainty in psychology are generally based on the assumption that the only relevant sources of uncertainty are those that cause random unsystematic variation in the conditional distribution of the dependent or response variable. This form of uncertainty is captured in the form of error terms that are assumed to be independently distributed with a mean of zero for any combination of values of whatever predictor variables are studied. This assumption is ubiquitous across many of psychologists' favoured statistical models, including regression and other linear models (Fox, 1997; Weisberg, 2005; Williams, Grajales, & Kurkiewicz, 2013), and linear mixed/multilevel models (Skrondal & Rabe-Hesketh, 2009). Practically, this assumption means that the effects of random sampling and/or random assignment to conditions, along with any uncorrelated random measurement error in the response variable, are accounted for as sources of uncertainty. But many other sources of uncertainty are ignored. In reality, claims about the impact of climate change on psychological variables are subject to a number of sources of uncertainty other than random sampling, random assignment, and random measurement error. In the following subsections, I will discuss some of these other sources of uncertainty.

5.2.3.1 Unacknowledged uncertainty in estimation: The problem of distributional assumptions.

Any application of inferential statistics—even via a non-parametric analysis—inevitably relies on distributional assumptions of some type, although the assumptions made can differ widely according to the analysis performed. Uncertainty about the validity of these assumptions is not taken into account in traditional inferential statistics such as confidence intervals and significance tests. The presence of added uncertainty surrounding estimates of relationships that exists due to uncertainty about whether the distributional assumptions made are actually met is rarely mentioned in published reports. In fact, researchers in psychology do not often check for assumption breaches in the first place (Hoekstra, Kiers, & Johnson, 2012). The added uncertainty surrounding estimates from statistical models implied by the presence of distributional assumptions has implications for the estimation of relationships, for causal inferences, and for predictions.

In some cases, the consequences of a breach of distributional assumptions can be relatively small, meaning that uncertainty about the validity of the assumption made increases the uncertainty of conclusions only slightly. For example, the assumption of normally distributed errors for linear models such as regression is not required for ordinary least squares estimation to produce coefficients that are unbiased, consistent, and efficient[18] (Wooldridge, 2009). Rather, normally distributed errors are required only in order to ensure that the sampling distribution of the coefficients is normal even if the sample size is small. Furthermore, the sampling distribution of coefficients will approach a normal distribution as the sample size grows larger even if the errors are not normally distributed, provided the errors are independently distributed. This means that significance tests and confidence intervals may still be trustworthy even without normally distributed errors. Gelman and Hill (2007) thus describe the normality assumption as "barely important at all" (p. 46) for regression models (see also Lumley,

[18] Broadly speaking, an unbiased estimator is one that produces estimates that, over repeated samplings, have the same expected value as the true population parameter. In other words, if we collected a very large number of samples, and calculated an estimate of a particular parameter in each sample using an unbiased estimator, the mean of the estimates would be the same as the population parameter. On the other hand, a consistent estimator is one that produces estimates that converge in probability toward the true parameter as the sample size grows larger. An efficient estimator is one that has less variance (is more accurate) than other choices of estimator. More formal definitions can be found in Ramachandran and Tsokos (2009).

Diehr, Emerson, & Chen, 2002). As such, uncertainty about whether error terms are actually normally distributed may not add dramatically to the uncertainty surrounding the substantive conclusions drawn from a statistical model.

On the other hand, other types of assumption breaches can have much more serious consequences. For example, consider the ubiquitous assumption that, conditional on the predictors, the expected value of the errors is zero. This is a common assumption that arises in many of the models most commonly used by psychologists, including the multiple linear regression model (R. A. Berk, 2004; Weisberg, 2005), related linear models such as the *t*-test, ANOVA, and ANCOVA, and linear mixed/multilevel models (Skrondal & Rabe-Hesketh, 2009). This assumption may be breached in several practical situations. For example, it is likely to be breached when random sampling is not used, when there is any measurement error in the predictor variables or measurement error in the response variable that is correlated with the predictors (Montgomery, Peck, & Vining, 2001), or when the form of a relationship is specified incorrectly such that unmodelled non-linearity is present. These issues are discussed in turn below.

5.2.3.1.1 Uncertainty due to the sampling method.

The assumption that the error terms have a conditional mean of zero in a statistical model (e.g., a regression model) is likely to be violated when random sampling is not used. Random sampling ensures that, for any given combination of the predictor values, the expected value of the errors over repeated samplings will be the same as the population mean error for this combination of predictor values. This mean will in turn be zero provided there are no problems with unmodelled non-linearity or measurement error. This assurance exists because the use of random sampling avoids any tendency of the sampling method to over-select cases that tend to have more positive (or more negative) errors. But without random sampling, no such assurance exists, and the errors may have conditional means that are different from zero (and that have different means at different levels of the predictors). This raises the problem of sampling bias. Importantly, the assumption that errors have conditional means of zero is required simply in order for the ordinary least squares estimator to be unbiased (Wooldridge, 2009). A breach of this assumption is thus much more serious than a lack of normality of the errors: It may lead to estimates that systematically over- or underestimate the true parameters. Of course, this implies in turn that confidence intervals and significance tests will not be trustworthy, as they do not take into account uncertainty arising due to the unknown quantity of bias. In general, the use of any sampling method other than simple random sampling with a perfect response rate means that uncertainty due to the use of a finite sample will not fully be captured in conventional measures of uncertainty such as significance tests and confidence intervals.

5.2.3.1.2 Uncertainty arising due to measurement error.

As alluded to above, problems with measurement error can result in uncertainty. This uncertainty arises because the distributional assumptions of commonly used statistical models will be breached if any measurement error is present (aside from purely random measurement error in the response variable). Specifically, commonly used statistical methods in psychology such as those based on the regression model assume that there is no measurement error in the predictor variables (Montgomery et al., 2001), and that any error in the measurements of the response variable is uncorrelated with the predictors (Williams, Grajales, & Kurkiewicz, 2013). These assumptions are implicit in the explicit assumption that the errors of prediction have a conditional mean of zero for any given combination of values on the predictors. In reality, however, it is probably rather unrealistic to assume a complete lack of any measurement error in the predictors, or of any correlated measurement error in the response. Breaches of these conditions can result in either upwardly or downwardly biased estimates of relationships (Williams, Grajales, & Kurkiewicz, 2013). This implies that uncertainty about the quantity and nature of measurement error (or in general, uncertainty about the validity and reliability of measurements) can contribute importantly to the quantity of uncertainty surrounding substantive conclusions.

5.2.3.1.3 Uncertainty about the form of relationships.

As mentioned above, the common assumption that the errors of a statistical model have a mean of zero for any combination of the predictor variables may be breached when the form of a relationship is incorrectly specified. Uncertainty about the form of relationships may thus imply uncertainty with respect to distributional assumptions, and therefore uncertainty with respect to substantive conclusions. Indeed, uncertainty about the form of relationships specified between variables is a problem that can increase the quantity of uncertainty surrounding conclusions about the nature and magnitude of effects that are already occurring, as well as the uncertainty surrounding predictions about the future impacts of climate change.

As an example of uncertainty about the true form of a relationship, some articles in the temperature-violence literature have suggested that the true form of the relationship between temperature and violent crime is linear (Anderson & Anderson, 1984; Bushman et al., 2005). Others have claimed that the relationship is non-linear, specifically in the form of an inverted U-shape (Bell & Fusco, 1989; Cohn & Rotton, 2005; Gamble & Hess, 2012; Rotton & Cohn, 2000). Claims of non-linearity are often linked to the negative affect escape model (Baron & Bell, 1975; Bell & Fusco, 1989). As discussed earlier in this thesis, the negative affect escape model suggests that moderate negative affect (as produced by moderately hot temperatures) can produce aggressive behaviour, but that very high levels of negative affect (perhaps produced by extremely hot temperatures) prompt escape. The negative affect escape model thus implies a relationship between temperature and violence that is not just non-linear, but also non-monotonic.

The method used to test this idea in almost all cases has been to specify a quadratic model and test the statistical significance of the squared term (Anderson & Anderson, 1984; Bell & Fusco, 1989; Bushman et al., 2005; Cohn & Rotton, 2005; Rotton & Cohn, 2000). In fact, this method is not particularly suitable for testing a hypothesis of non-monotonicity. In a quadratic model such as $\hat{Y} = \beta_0 + \beta_1 X + \beta_2 X^2$, evidence of a non-zero slope β_2 for the squared term *can* be indicative of a non-monotonic relationship (such as the reverse-U shaped relationship implied by the negative affect escape model), but that is not necessarily the case. The squared term can have a non-zero slope even if the true relationship is non-linear but still monotonic (Lind & Mehlum, 2010). Of course, at some point the relationship between temperature and interpersonal violence *must* become negative, since at extremely high temperatures humans will be unable to survive, let alone assault one another (as acknowledged by Anderson, 1989). As such, the relationship between temperature and violence cannot be monotonic across *all* possible temperatures; rather, the question is whether it is monotonic within the range of outdoor temperatures actually observed in everyday life. The problem is that the method that has been used to try and determine the true shape of the causal relationship within the range of temperatures observed is unsuitable for that purpose.

Inappropriate statistical methods can contribute to uncertainty surrounding the true form of relationships, but even without such problems some uncertainty about the form of relationships inevitably remains. This is especially the case in psychology, where the lack of mathematical theory (see Myung & Pitt, 2002) means that decisions about the form of relationships may often need to be based on observations alone. This is not to say that there are no useful tools to help with this process. For example, fit statistics that attempt to balance parsimony and model fit such as the Akaike Information Criterion may be used to select between competing models (Hurvich, Simonoff, & Tsai, 1998). Nevertheless, it will always be difficult to completely rule out any uncertainty about the true form of any given relationship.

Obviously, uncertainty about the true form of relationships between variables is a problem in psychological research generally, and not just psychological research concerned with the impacts of climate change. However, when studying the impacts of climate change, uncertainty about the true form of relationships has an important flow-on effect: It produces quantitative uncertainty about the effect of a climatic change of a given magnitude (e.g., the effect a temperature increase of 1°C on assaults). Models with different functional forms (linear, log-linear, quadratic, spline, etc.) would imply different conclusions about the size of such an effect.

5.2.3.2 Uncertainty in causal inference.

One purpose for which statistical inference can be used is to make conclusions about variables and relationships in a population based only on a sample of data from that population. Importantly, conventional descriptions of the assumptions of statistical analyses refer to the assumptions necessary to achieve this restricted kind of inference: that is, inferences about variables and relationships in a population, *not* inferences about causal effects. For example, to achieve trustworthy estimates of the relationships between a set of predictors and a response variable in a population using multiple regression one need only assume that the model error terms are independent, homoscedastic, normally distributed, and have conditional means of zero (Williams, Grajales, & Kurkiewicz, 2013). However, a fulfillment of these assumptions is necessary but not sufficient to justify treating the estimated relationships as representing causal effects. This is a concern that is obviously relevant to studies concerned with estimating the effects of climate change on behavioural variables based on observed data. However, it is also relevant when using estimates of causal effects to predict behaviour in the future. Furthermore, it is a concern that arises not just when using sample data, but also when analysing data from entire populations. Even if the use of data from an entire population means that statistical relationships can be described without uncertainty (at least for the population and time period studied), inferences about causal effects will remain subject to uncertainty. This is one reason why inferential analyses may be required even when using data from an entire population, as was the case for many of the empirical analyses reported in this thesis.

In discussing causal effects here I will use the *potential outcomes* framework (Holland & Rubin, 1988; Rubin, 2005). For simplicity I will consider the example of estimating the causal effect of a single "treatment" to which units are or are not exposed. However, the framework can be extended to estimating the causal effects of differences in continuous variables (such as temperature). In the potential outcomes framework, the causal effect of a treatment on the dependent variable for a unit (e.g., a person) i can be defined as the difference[19] between the potential outcomes $Y_i(0)$ and $Y_i(1)$. $Y_i(0)$ is the value of the dependent variable that unit i would obtain if assigned to the control condition, and $Y_i(1)$ is the value the unit would receive if assigned to the treatment. Obviously, we can observe only one of the potential outcomes for any given unit. This is "the fundamental problem of causal inference" (Holland, 1986, p. 947). We can nevertheless obtain an estimate of the average treatment effect $E(Y_i(1) – Y_i(0))$ for a sample of units using observed data. This data may or may not be based on a randomised experiment, and the average treatment effect may or may not be estimated after conditioning on a set of covariates (i.e., control variables). But for the resulting estimate of the causal effect to be unbiased, a specific set of assumptions do need to hold:

1) The stable unit treatment value assumption, which has two parts (Rubin, 2005):
   a) Whether or not a particular unit (e.g., a person, a time period, etc.) is subjected to a treatment does not influence the potential outcomes of other units
   b) All units exposed to a particular condition receive the same treatment.

[19] It is also possible to define a treatment effect in ways other than as a simple difference between potential outcomes. For example, one might use the *ratio* of the potential outcome in the treatment condition to the potential outcome in the control condition (Rubin, 2005).

2) The strong ignorability assumption: Conditional on any included covariates *X*, whether or not a unit receives treatment is independent of the potential outcomes (Rosenbaum & Rubin, 1983). In other words, how a unit would respond to being in the treatment condition, and how it would respond to being in the control condition, are both independent of which condition the unit was *in fact* assigned to. This assumption would be breached if an omitted confounding variable was present, or if the supposed dependent variable in fact affected the independent variable (i.e., if the causal relationship was in fact in the opposite direction to that hypothesised).

Uncertainty about whether these conditions are met will obviously be a source of uncertainty with respect to any causal inferences made about the effects of climate change. The conditions specified may be plausible in a randomised experiment: For example, use of a randomised experiment implies strong ignorability (Rosenbaum & Rubin, 1983). But psychologists studying climate change impacts often use observational rather than experimental data. Indeed, that was the case for the empirical studies reported in this thesis. With observational data, the conditions above may not be plausible. For example, consider temporal studies of the relationship between temperature and violence, in which the units studied are time periods (such as dates). In such studies, the stable unit treatment value assumption might be breached if there are lagged effects of temperature: That is, if the fact that a particular date had a hot temperature results in an increased risk of violence on the *next* day. Similarly, the strong ignorability assumption may not always hold: The possibility of uncontrolled confounders will in most cases be difficult to rule out with confidence.

Psychologists are generally well aware of the risk of confounding when attempting to draw conclusions from non-experimental designs. Certainly this issue was discussed explicitly in the three empirical articles presented in this thesis. However, the key point is that when uncertainty with respect to the size of some effect is depicted in the form of an interval estimate such as a confidence interval, this interval estimate will in almost all cases *not* account for uncertainty arising due to the possibility of uncontrolled confounders (or due to the possibility of a breach of the stable unit treatment value assumption). Thus, if interpreted as measures of uncertainty with respect to the size of causal effects, such interval estimates underestimate the magnitude of uncertainty present.

5.2.3.2.1 Uncertainty in causal inference due to varying effects.

There is an additional and less well recognised problem when attempting to make causal inferences from observational data: that of varying treatment effects. When using observational data, the use of covariates (control variables) is generally crucial in order to make the strong ignorability assumption at all plausible. Although the stable unit treatment value and strong ignorability assumptions are sufficient to obtain an unbiased estimate of the average treatment effect in principle, this relies on the researcher being able to directly calculate the mean difference between the dependent variable observations for units in the treatment and control conditions, *conditional on the values of the covariates*. In reality, of course, this process of conditioning —of removing the effects of control variables—must be achieved by estimating some sort of statistical model. Often this might be achieved using a method such as multiple linear regression estimated using ordinary least squares. But ordinary least squares estimation does not take place by literally calculating an average difference between treated and control units after conditioning on covariates. Rather, the ordinary least squares estimation procedure minimises the sum of squared differences between the predicted and observed values of the dependent variable.

As it happens, if the effect of the treatment is the *same* for all units, and the distributional assumptions of the regression model hold, along with the stable unit treatment value and strong ignorability assumptions, then the regression model will provide an accurate estimate of the average treatment effect. But this is not the case if the effects of treatment vary. Aronow and Samii (in press) point out this problem, noting that a regression model[20] fitted using ordinary least squares does not weight the contributions of each unit equally. Specifically, they demonstrate that the model will give more weight to those units whose values on the independent variable are not explained well by their values on the covariates. If treatment effects vary, this unequal weighting means that the ordinary least squares estimate for the slope of the independent variable will not be a representative estimate of the average treatment effect. The possibility of varying treatment effects is thus another source of uncertainty in drawing causal inferences. It is a source of uncertainty that is particularly relevant to

[20] Aronow and Samii (in press) also show that this problem applies also to non-linear models estimated using maximum likelihood, including generalised linear models, and to models including random coefficients, such as mixed models.

the study of climate change impacts, where it seems *a priori* implausible that the impacts of particular climatic changes will be constant across people, places, and times.

The problem of varying treatment effects means that even when the distributional assumptions of a model are met, and the additional assumptions necessary to draw causal inferences are also met, the possibility of effects that vary across units may still be a source of additional uncertainty with respect to an estimate of the average effect of some variable. The implications of varying treatment effects have obtained some attention in the social sciences, and particularly political science (Gelman, 2004, 2015; Gelman & Huang, 2008). However, psychologists typically have little to say about whether effects are variable or constant (Sohn, 1999), and are likely to be unaware of the implications of varying effects for drawing causal inferences.

Varying effects also have implications for the validity of using an estimated effect based on observations from a broad set of units to make conclusions about an effect on an *individual* unit. For example, even if one knew without any uncertainty the average effect of a 1°C increase in temperature on assault in New Zealand, there would *still* be uncertainty about the size of the effect of a 1°C increase in temperatures within a specific location (e.g., Auckland). The effect of such a temperature increase could vary across times, places, and populations, and in fact seems likely to do so.

5.2.3.3 Uncertainty in prediction.

Thus far, I have discussed sources of uncertainty with respect to inferences about relationships and causal effects. But there is another type of inference of particular importance to researchers interested in the potential impacts of climate change: inferences about the future. While many of the sources of uncertainty discussed above are relevant when attempting to make predictions, there are some additional sources of uncertainty that are specific to the making of predictions about the future. These sources of uncertainty are rarely modelled or addressed in psychological research. Importantly, the fact that researchers may wish to draw conclusions about phenomena or effects occurring in the future is another reason why the reporting of inferential statistics and information about uncertainty may be necessary even when analysing data from an entire population (as was the case for several of the empirical analyses reported in this thesis). Conceptually, data about the behaviour of a population over one period of time can be thought of as a sample of the behaviour of that population over a longer period of time. However, the behaviour of a population at one point in time may not be representative of its behaviour at other points in time

5.2.3.3.1 Uncertainty about whether model parameters will change over time.

In climatology, predictions about the future climate are based in large part on knowledge about chemical reactions and atmospheric processes that are well described by theory. For example, as mentioned in section 5.2.2.1.1, the general circulation models that are used to predict future changes to the Earth's climate are based on the Navier-Stokes equations, which represent Newton's second law of motion applied to the motion of fluids. The theory applied by climatologists is presumed to apply reasonably accurately, both in the past and in the future. There is some legitimate justification for this belief: The theory instantiated in the Navier-Stokes equations, for example, is a general theory of the motion of objects (Newton's laws of motion), and the accuracy, scope and limitations of this theory are well-established (Skinner, 1982).

On the other hand, in psychological climate change research, inferences about the future will typically be based on statistical models with parameters estimated based on data from the past. The relationships estimated by statistical models are not theoretically produced universal laws; they are observations about relationships that occurred in a particular time and context. These relationships may change in the future. For example, even if the estimates of the effect of temperature on assault, suicide, and self-harm hospitalisations produced in the empirical section of this thesis are perfectly valid for the period of time studied—itself a claim that would obviously be subject to uncertainty—these causal relationships may change in the future. One cause of such a change might be an interaction between temperature and some other variable that changes over time. Obviously it is impossible to rule out such a change in the value of parameters: In fact, this issue is just a special case of the problem of induction (see Sloman & Lagnado, 2005). In a way, this source of uncertainty might be considered part of the general problem of varying treatment effects, as discussed previously. However, this is a particularly vexing problem in the case of predictions about the future. Even if we knew the (current) average effect for some variable in a population, and how this effect varied across particular places and people, there would still be the possibility that this effect could change in the future.

5.2.3.3.2 Uncertainty about causal inputs.

Climatologists making predictions and projections about the future climate do so based on a range of different emissions scenarios, representing different assumptions about how anthropogenic greenhouse gas emissions will change in future (see Moss et al., 2008). Similarly, predictions about human behaviour in a warming world are likely to involve assumptions about causal inputs[21] such as future changes to the Earth's temperature or other climatic variables. This is likely to be the case whether the predictions sought are direct predictions about actual future behaviour or conditional predictions about the *impact* of climate change. The future values of the causal inputs that these predictions will be based on—e.g., climatic variables—will be subject to significant uncertainty.

A very small number of studies in the temperature-violence area have made predictions about the effects of a range of different temperature increases. For example, Anderson (2001) reported estimated increases in annual violent crime in the US for temperature increases from 0 to 8°F. However, Anderson left it to the reader to judge the relative plausibility of the different values contained within this wide range of temperature change. This meant that the magnitude of uncertainty about future temperature changes (and the associated uncertainty about the impact of these temperature changes on the incidence of violence) was not effectively communicated.

5.2.3.3.3 Uncertainty due to stochastic behaviour of the predicted variables.

Predictions about the impact of climate change on psychological variables will nearly always need to be made on the basis of statistical models that include a stochastic (probabilistic) component. This stochastic component often takes the form of a random error term in a statistical model, as discussed previously. Importantly, the presence of stochastic uncertainty means that even if the form, parameters and causal inputs for a model were known with absolute certainty, with no possibility of changes to parameters in the future, there would *still* be some uncertainty about actual future observations of the behaviour predicted by the model.

When making predictions about future observations, *prediction* intervals can be used, which directly indicate the magnitude of uncertainty with respect to the values of

[21] Of course, this is not *necessarily* the case: For example, one might generate predictions about the future values of a variable based purely on past observations of that variable. This method seems likely to produce predictions of very limited accuracy, however, especially over long timeframes.

actual future observations (Chatfield, 1993). I know of no studies concerned with the temperature-violence relationship that have made use of prediction intervals. This said, uncertainty occurring due to the stochastic nature of behavioural variables is a problem primarily when attempting to make predictions about actual future events such as the number of assaults occurring within a particular year. It does not need to be given especial attention when attempting to predict the long-run *effect* of a temperature increase of a given magnitude. The predictions made about interpersonal violence in the empirical studies reported in this thesis were of the latter type. So too were the predictions about climate change and violence reported in other studies discussed in this thesis (e.g., Anderson, 2001; Anderson et al., 2000; Anderson & DeLisi, 2011; Rotton & Cohn, 2003).

5.2.3.4 Reporting uncertainty.

Setting aside the issue of what sources of uncertainty are accounted for as part of a study's design and data analysis, there remains the task of accurately and clearly *communicating* that uncertainty. Unfortunately, psychologists' conventional methods of statistical inference have significant limitations as tools for reporting uncertainty.

5.2.3.4.1 Significance tests.

In psychology, the most common method used to report quantitative uncertainty is to report the results of a null hypothesis significance test. A study by Cumming and colleagues (2007) found that around 97% of articles in leading empirical journals in psychology report the results of one or more significance tests. Significance testing is based on the frequentist interpretation of probability, within which the *probability* of an event refers to the limit of the proportion of times that the event occurs as the number of trials increases toward infinity (Lyons, 2013). For example, a frequentist can say that the probability that a tossed coin will come up heads is .5, because if we tossed the coin a large number of times, the proportion of heads would converge toward .5. A frequentist stance thus permits probability statements only about events when their frequency can (at least hypothetically) be tallied over many trials. The frequentist interpretation of probability does *not* allow us to make any useful statements about the probability of hypotheses about the world that are simply either true or false. Because the frequentist interpretation of probability allows for conclusions about the probability of only some types of statements, frequentist statistical analyses (and especially significance tests) can only provide probability estimates that, in many cases, relate only indirectly to the hypotheses that researchers actually wish to test.

The fact that frequentist analysis cannot always provide clear and intuitive statements about uncertainty is made clear when we consider the correct interpretation of a *p* value. A *p* value indicates the probability of observing a test statistic as or more extreme than that observed, if the null hypothesis is true (Goodman, 2008). This counter-intuitive interpretation means that the reporting of *p* values does not facilitate clear communications about uncertainty. The *p* value is a measure of uncertainty, in an indirect sense: All else being equal, the greater the uncertainty surrounding an estimate, the greater the *p* value. However, the *p* value is not a measure of uncertainty that directly answers the questions likely to be of most interest: How certain are we that the null hypothesis is false? What is the probability that the parameter is in a particular direction? And perhaps most importantly, what is the interval of values within which we can be confident that the parameter actually *does* fall?

There is a very substantial body of literature concerned with the limitations of null hypothesis significance testing. For example, the vast majority of applications of significance testing specify a null hypothesis that a particular parameter or relationship has a value of exactly zero, even though such a "nil" hypothesis is almost always implausible (J. Cohen, 1994, p. 1000). Furthermore, the correct interpretation of a *p* value can change depending on the subjective intentions of the researcher (e.g., when optional stopping is used in clinical trials: Wagenmakers, 2007). Even more worryingly, the use of significance testing and the accompanying preference of authors, reviewers and editors for studies producing *p* values below the arbitrary 0.05 threshold leads to a systematic bias toward significant findings in the published literature (Easterbrook, Gopalan, Berlin, & Matthews, 1991). These problems notwithstanding, for present purposes it is enough to recognise that null hypothesis significance tests simply do not directly answer the questions likely to be of most interest to researchers and users of research. This is especially the case in the climate change arena, when the *size* of effects and future changes is of primary concern.

Importantly, because significance tests do not provide clear and direct information about the magnitude of uncertainty surrounding effect size estimates, they are extremely limited as a method for reporting uncertainty. This is the case even when they are accompanied by a point estimate of standardised or unstandardised effect size. In the temperature-violence literature, it is rare for a study to report only significance tests without any indication of effect size (be it standardised or unstandardised). However, it is common for studies in this area to report point estimates indicating the size of effects, but include only significance tests as indicators of the uncertainty surrounding these effects (e.g., Ajdacic-Gross et al., 2007; Ceccato, 2005; Cohn, Rotton, Peterson, & Tarr, 2004; Lester, 1999; Oravecz, Czigler, & Moore, 2006; Preti et al., 2007). This means that while the best estimate of the population parameter is communicated, and so too the probability of observing an estimate that large or larger if the true parameter was zero, no information is provided about the range of values within which the true parameter is most likely to fall. The reader is left unsure whether the true population parameter is likely to fall in a narrow interval very close to the sample estimate or whether it could actually plausibly fall anywhere within a wide range of values. Without being provided an interval estimate, readers with limited statistical expertise may even assume that there is *no* uncertainty about the size of the true parameter. Obviously, then, the sole use of significance testing is unsuitable as a method for reporting uncertainty in psychological research concerned with climate change (whether it is ever a suitable method for reporting uncertainty is a controversial question; see Nickerson, 2000).

5.2.3.4.2 Confidence intervals.

The other commonly used method for reporting uncertainty in psychology is the confidence interval. Confidence intervals are potentially a more useful tool for describing the uncertainty surrounding conclusions relating to the potential impacts of climate change because they provide quantitative information about the uncertainty surrounding the size of a relationship or effect. Confidence intervals were used heavily in all three empirical studies in this thesis. This meant that information about effect size was reported not just in the form of the best point estimate, but also in the form of an interval within which we could be “confident” that the true parameter actually falls. Confidence intervals have also been used extensively in other studies concerned with the relationship between temperature and violence (some of many examples include Anderson et al., 2000; Dixon et al., 2014; Helama et al., 2013; Hsiang et al., 2013; Rotton & Cohn, 2003).

But what does it mean to be confident? Confidence intervals, like significance tests, are located within the frequentist tradition and thus provide information about uncertainty that only indirectly answers the questions that are likely of most interest to the user. Strictly speaking, a 95% confidence interval for a parameter can be interpreted as follows: If the study was replicated a very large number of times, and a 95% confidence interval for the parameter calculated every time, then 95% of the resulting intervals would include the true parameter (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2015; Neyman, 1937). This probability statement refers only to what we expect to happen over multiple trials: Technically speaking, the 95% confidence interval calculated in an *individual* study has no clear interpretation. When guided by the frequentist interpretation of probability, we cannot say that there is a 95% probability that a *specific* calculated confidence interval contains the parameter of interest. The parameter is a fixed value, not something that changes over repeated trials. By a frequentist definition of probability, the probability that the true parameter falls within a specific calculated interval is either one, if the parameter does fall within the interval, or zero, if it does not. The fact that a specific confidence interval lacks a clear and intuitive interpretation is probably one reason why many psychological researchers misinterpret confidence intervals (see Hoekstra, Morey, Rouder, & Wagenmakers, 2014). It also means that confidence intervals may not necessarily be an ideal way to communicate uncertainty about climate change impacts. The topic of whether there are better ways to calculate and report interval estimates of uncertainty, and whether more intuitive interpretations of confidence intervals can be justified, is returned to in the recommendations chapter that follows (see section 6.5.1.2).

5.3 Conclusion

In sum, there are several important problems with psychologists’ favoured methodological strategies that may cause difficulties when attempting to produce effective and useful research concerned with the impacts of climate change. These problems include measurement and analysis strategies that hamper the clear communication of the sizes of effects; a neglect of the role of time; a lack of theory development that facilitates prediction-making; and the use of inadequate strategies for modelling and reporting uncertainty. If left unaddressed, these problems may seriously hamper psychologists’ capacity to produce useful research pertaining to either the past or future impacts of climate change. Fortunately, however, there is a range of strategies that psychologists can use to deal with these problems. In the following chapter, a number of recommendations for dealing with the problems identified in this methodological critique are provided.

6 Methodological Recommendations

This final chapter provides several recommendations for psychological researchers aiming to study the impacts of climate change. The recommendations offered are not intended to comprise a comprehensive guide for conducting research concerned with the psychological impacts of climate change. Rather, they are offered as potential solutions to the specific problems identified in the methodological critique.

6.1 Use Measurement Strategies that Facilitate Clear Reporting of Effect Size

In the methodological critique, I argued that psychological research projects concerned with the impacts of climate change need to clearly report the size of climate change effects. Clearly reporting effect size is not just a matter of the choice of statistical analysis. Rather, decisions made early in the research process—particularly with respect to measurement strategies—can have an important impact on how clearly the size of effects can be reported. In the methodological critique, I identified measurement strategies that can hinder the clear reporting of effect size. These strategies include unnecessary transformations of quantitative variables and the use of psychometric measurement scales without clear units of measurement.

Some measurement strategies that are problematic for clear reporting of effect size can easily be avoided. For example, researchers using meteorological measurements that have quantitative properties and clearly defined units of measurement (e.g., temperature in °C, rainfall in mm, etc.) should avoid applying transformations, summations, or categorisations to these variables that result in observations that lack clear units of measurement (e.g., as in Anderson & Anderson, 1996). When a physical quantity is used as a predictor or explanatory variable, using a form of the variable with clear units of measurement can help to ensure that effect size is reported clearly. As one example, when temperature is used as a predictor of some behavioural variable, the fact that effect size can be reported in terms of the magnitude of behaviour change per degree Celsius can facilitate the clear reporting of effect size. This is provided, of course, that the behavioural variable of interest can also be recorded on a scale with clear units of measurement.

Recording human behaviour on scales with clear units of measurement may indeed represent a more significant challenge. Psychologists frequently make use of scales that are intended to have quantitative properties (e.g., a summed score on a rating scale, or a test of cognitive ability). However, the units on such scales rarely have a clear interpretation, and the meaning of a one-unit difference may even vary across different points on the scale: Psychological measurement devices typically produce (at best) ordinal scores. As mentioned previously in section 5.2.1.1, there is no evidence that psychological attributes possess the quantitative properties necessary to permit interval or ratio measurement (Michell, 2008, 2009, 2012). Again as discussed earlier, the use of ordinal measurements of behaviour makes it difficult to effectively communicate effect size: If the units on the measurement scale used have no clear and constant meaning, then effect size cannot effectively be communicated in terms of the number of units by which a behavioural variable changes given a particular change in conditions.

It is of course possible that some psychological attributes may be genuinely quantitative. If this is the case, then it should be possible to develop scales to measure these attributes that have units with a clear and constant meaning. Some authors (e.g., Kyngdon, 2011; Michell, 1997) suggest that it may be possible to use statistical analysis (specifically the conjoint measurement model; Luce & Tukey, 1964) to test a hypothesis that a particular psychological attribute is quantitative. On the other hand, Trendler (2009) argues that it is fundamentally impossible to obtain the necessary experimental control over psychological attributes to satisfactorily test this proposition.

I would argue even further that while it is possible that there exist variables that are distinctly psychological and also quantitative, there is little *a priori* reason to assume that this is the case. In the physical sciences, only a very small number of independent quantities have been identified. In fact, the International System of Units (SI; the most well-known measurement system) has just seven base units: the metre for length, the kilogram for mass, the second for time, the candela for luminous intensity, the ampere for electrical current, the kelvin for temperature, and the mole for the amount of a substance (see International Bureau of Weights and Measures, 2006). All other SI units are derived from and dependent on these base units. For example, the unit of *force* is the Newton, a unit derived from the base units of kilogram, metre, and second: One Newton is defined as the force necessary to accelerate a mass of 1 kg at the rate of 1 metre per second per second. Scientists in a variety of different fields—from physics to biology and ecology—use this same system of SI units, derived from a small number of known independent quantities. It is not the case, for example, that biologists have identified unique biological quantities: Rather, biologists measure the same physical quantities as other scientists, just applying these measurements to variables of particular biological interest (e.g., distance of a migration, mass of an animal, etc.). In other words, it appears to be the case that quantitative variables are scarce in the real world. Only a very small number of variables that both have quantitative properties and that also cannot be represented simply as derivations of other quantitative variables have ever been identified. Of course, it may be the case that psychologists ultimately discover that human behaviours are determined by physical quantities, such as the mass of particular neural substrates. However, the idea that there exist variables that are both uniquely psychological (i.e., *not* derivable from known physical quantities) and that also have true quantitative properties appears to be rather implausible. If psychologists continue to attempt to use psychometric scales that treat psychological attributes as quantitative, but that do not actually produce quantitative measurements with clear units, then it will be continue to be difficult to clearly communicate the size of effects. There may be some strategies that could help improve the usefulness of traditional psychometric measurements, but psychologists may also need to consider new and unfamiliar measurement strategies.

6.1.1 If using psychometric scales, use contemporary methods for investigating reliability, validity, and quantitativity.

If psychological attributes are not quantitative in the first place, then the goal of identifying a single clear unit of measurement with the same meaning across different levels of any psychological attribute is doomed to begin with. Furthermore, even to the very limited extent that any psychological measurement device has a meaningful unit of measurement, the unit of measurement will inevitably be unique to that device. This is a practical problem even for researchers who are not concerned with the abstract question of whether psychological attributes are quantitative in the strict mathematical sense. Researchers might wish to justify the use of numeric psychometric scales on the pragmatic grounds that such scales are *useful*, but this seems an uncomfortable position: Psychometric scales have some important uses, but the fact that the application of such scales prevents the reporting of effect size using a unit of measurement with a clear and constant meaning implies a major impediment to their usefulness in many situations. Certainly this would be a major and important restriction to the usefulness of such scales for researchers wishing to communicate information about the size of climate change impacts.

Nevertheless, it seems realistic to expect that many psychologists may wish to continue using conventional psychometric measurement scales. Several broad strategies may to at least some degree improve the usefulness and justifiability of this approach. One strategy would be to clearly outline how the units of measurement of the scales used can be interpreted, and correspondingly how the estimates of effect size produced can be interpreted. As alluded to above, another strategy would be to consider using the fundamental conjoint measurement model to determine whether the measurement scales used in a particular project are genuinely quantitative (see Borsboom, 2005; Kyngdon, 2011; Luce & Tukey, 1964).

It could also be useful to attend to recent developments in psychometrics with respect to reliability and validity. For example, Sijtsma (2008) pointed out several problems with the ubiquitous Cronbach's alpha measure of reliability, showing that alpha is neither equal to a test's reliability in realistic situations, nor a useful indicator of internal consistency (i.e., of whether or not the scale is unidimensional). Alternative contemporary measures of internal consistency reliability that more clearly reflect the internal structure of a test are available (see Revelle & Zinbarg, 2009; Zinbarg, 2006), but rarely used. With respect to validity, an important contemporary development that researchers should be aware of if seeking to justify the use of psychometric scales is Borsboom, Mellenbergh, and van Heerden's (2004) redefinition of validity (cf. Cronbach & Meehl, 1955; Messick, 1995). Borsboom et al. argue that a test is a valid measure of an attribute if and only if the attribute exists, and variation in the attribute has a causal effect on scores on the test. This intuitive definition of validity implies the need for a different approach to test validation than conventional strategies such as correlations between tests and criterion measures. Rather, Borsboom et al. suggest that a crucial step in the validation process is the generation of a hypothesis that outlines the *process* via which variation in the attribute causes variation in test scores. Only once this hypothesis is specified can empirical tests be used to determine whether the process specified actually occurs.

In sum, while realistically psychologists studying climate change impacts may at least in some cases wish to continue using traditional psychometric scales, they—and all psychologists—need to consider carefully about how they will justify the assumptions that their measures are quantitative, reliable, and valid. These concerns about measurement apply especially when studying the impacts of climate change, a context within which there is a particular need for clear statements about the magnitudes of effects and predicted changes to behaviour.

6.1.2 Consider categorising behaviour rather than attempting to measure psychological attributes.

Given the problems with purportedly quantitative psychometric scales raised above, there may be a need for more radical changes to measurement practice. An effective alternative to using psychometric scales and treating psychological attributes as quantitative might be to develop and use clearly defined *categorisations* of behaviour or individuals. This approach was used extensively in the empirical section of this thesis. I did not attempt to use quantitative scales attempting to measure the degree to which participants were inclined to suicidality, self-harm or aggression. Rather, the behavioural variables studied were *counts* of suicide, self-harm, and assault. Relatively clear definitions of what constituted an act of suicide, self-harm, or assault were applied. Along with the use of genuinely quantitative variables with clear units of measurement as predictors (especially temperature in °C), this meant that the effect sizes reported in the empirical studies had clear and interpretable units of measurement. In most cases, the effect sizes reported took the form of the proportional change in the frequency of a behaviour (e.g., assault) given a 1°C increase in temperature.

Conventional wisdom has it that using continuous measurements (or at least scales with many possible values) is preferable to using simple categorisations because the added variability in responses on a continuous scale improves statistical estimates of reliability and validity (MacCallum, Zhang, Preacher, & Rucker, 2002; Markon, Chmielewski, & Miller, 2011), and improves statistical power (Altman & Royston, 2006; J. Cohen, 1983; MacCallum et al., 2002). These considerations may partially motivate psychologists' preference for purportedly quantitative scales of psychological attributes, and a general lack of interest in developing taxonomies of human behaviour (but see Michell, 2011 for a discussion of historical reasons for psychologists' preference for attempting to quantify psychological attributes)[22].

[22] The primary exception to psychologists' general preference for quantification rather than classification of behaviour is in clinical psychology, where the diagnostic and statistical manual of the American Psychiatric Association (2013) is used to define mental disorders as categorical entities rather than treating each disorder as falling on a continuum.

Of course, the reliability and validity of measurements do matter when performing psychological research concerned with climate change impacts, as in any research setting. However, of even more importance is the meaningfulness, validity and usefulness of *research findings*. When a research finding takes the form of an estimate of effect size, and that effect size estimate is couched in terms of a unit of measurement that has no clear and constant meaning, then it is difficult to see how the finding could be valid or useful. Statistical power likewise does matter, but there seems to be little point in achieving high power to detect an effect if the size of that effect cannot be clearly communicated.

On the other hand, when behaviours are clearly defined and categorised, the clear communication of effect size is facilitated. Research concerned with the effect of exposure to natural disasters on mental health is one area relevant to climate change in which some researchers have applied this strategy fruitfully. For example, McFarlane and Van Hooff (2009) found that 21.5% of 529 Australians who had lived in a part of the country devastated by the 1983 Ash Wednesday fires experienced an anxiety disorder (other than post-traumatic stress disorder [23]) at some point over a 20-year follow-up period, in comparison to just 16% of controls. In other words, McFarlane and Van Hooff's dependent variable was a categorical variable simply indicating whether or not a participant experienced an anxiety disorder or not. The resulting findings reported in the form of percentages (and the implied 34% percentage increase in anxiety disorder risk associated with bushfire exposure) give a much clearer indication of the size of the effect of exposure to this natural disaster than would a mean difference on a numeric psychopathology scale (e.g., as in the study of the effect of exposure to disaster on mental health by Baum et al., 1992).

Of course, classifying behaviour into clearly defined categories is not always easy. In the empirical articles reported in this thesis, the classifications of behaviour used were not so much developed as imposed from without: The Ministry of Health (and the International Classification of Diseases; see Centre for Disease Control and Prevention, 2013) defined what constituted an act of suicide, or an act of self-harm that contributed to a hospitalisation, and the police (and New Zealand law) defined what constituted an assault. Thus, relatively little thought was required to apply these categorisations, beyond dealing with practical issues such as the presence of multiple

[23] Interestingly, McFarlane and Van Hooff (2009) did *not* find a statistically significant relationship between bushfire exposure and the incidence of post-traumatic stress disorder.

hospitalisations for the same act of self-harm and so forth. In other contexts, psychologists might need to more carefully decide what criteria will be used to determine whether a person or behaviour falls into a particular category. For example, measuring the effect of exposure to an extreme weather event on the risk of a psychological disorder requires carefully specifying what criteria will be used to determine presence of that psychological disorder in the particular research project. Some options might be structured observations, a clinical interview, or a classification based on responses to a self-report scale.

In other cases researchers might need to not only decide what criteria will be used to determine membership to categories, but also what categories will be used in the first place. For example, psychologists might well be interested in the impact of climate change on subjective well-being. Subjective well-being is traditionally assessed using psychometric scales (see Pavot, 2008). But such scales lack meaningful units that can be used to usefully report climate change effects. Therefore, researchers might need to produce a classification of well-being that includes clearly defined and meaningful categories of well-being (or lack thereof). The impact of climate change on well-being could then be reported in terms of how it changes the distribution of individuals falling into different categories of well-being.

The strategy of using categorisation rather than traditional psychometric measurement presents challenges, but may be valuable. What I am arguing for is a shift away from assuming that psychological attributes exist in the world (independently of our activities as psychologists[24]), with the researcher's task being simply to find devices that satisfactorily measure these attributes. Rather, psychologists may instead need to *decide* how best to classify and categorise human behaviour. The fact that decisions are required in terms of what categorisations are used should not be regarded as lamentable. Biologists, for example, accept that classification systems are necessarily theory-dependent, rather than reflections of naturally existing divisions and orders, but nevertheless see classification as a crucial scientific task (see Mayr & Bock, 2002). Psychologists might benefit from using biology as their model in this respect, rather than trying to emulate physicists, who focus on measuring dimensions rather than

[24] Psychologists implicitly take this *realist* ontological stance with respect to psychological attributes when they treat psychological attributes as latent variables in common statistical analyses such as exploratory factor analysis and structural equation modelling (Borsboom, Mellenbergh, & van Heerden, 2003).

classifying objects. To the extent that psychologists do currently discuss classification and taxonomy, the focus is often and regrettably restricted to taxonomies of purportedly quantitative traits, such as the Big Five trait taxonomy (see John, Naumann, & Soto, 2008).

6.1.2.1 Assessing the reliability and validity of categorisations.

If psychologists studying climate change impacts attempt to categorise rather than measure behaviours and attributes, they will still need to consider the reliability and validity of the resulting categorisations. When behaviours or individuals are categorised by raters or decision makers, the interrater reliability of these categorisations can be examined using statistics such as Cohen’s kappa or Fleiss’ kappa for unordered/nominal categorisations, or Kendall’s *W* for categorisations on an ordered scale (Gisev, Bell, & Chen, 2013).

Borsboom and colleagues’ (2004) definition of validity can also be a useful framework for considering validity even when the “measurements” made are actually categorisations of behaviour or individuals. As discussed previously, this definition of validity suggests that a test is a valid measure of an attribute if (1) that attribute actually exists, and (2) if variation in the attribute causes variation in scores on the measure. This definition can be translated to the context of investigating the validity of acts of categorisation rather than measurement.

At first glance, condition (1) stated above might not seem to be particularly relevant in this context: When researchers use categorisations, there is no assumption necessary that some abstract quantitative attribute exists to be measured. This said, there is still an assumption necessary that the characteristics that the units will be categorised with respect to actually exist. It is possible to imagine scenarios in which this requirement might not be met. A categorisation of individuals by aura colour, for example, might fail this test, since there is no evidence that psychic auras actually exist (Perez, 2011). Outside of such rather outlandish scenarios, however, it seems likely that the categorisations that psychologists might use will involve characteristics (such as the performing of specific behaviours) whose existence need not be subject to be much debate. There is little question that some individuals actually assault each other, commit suicide, or perform acts of self-harm, for example.

Condition (2) for validity—that variation in the attribute causes variation in measurement outcomes—remains directly relevant in the context of categorisation, however. Condition (2) for validity can usefully be restated in the context of a decision of categorisation as such: If a unit (behaviour or individual) *X* actually belongs to category *Y* then this increases the probability that unit *X* will be categorised as belonging to category *Y*. Researchers using categorisations of behaviour to study climate change impacts will need to consider and specify how this causal process takes place. For example, with respect to Study Three, we can hypothesise that performing an act of self-harm increases the probability that an individual will be admitted to hospital, and, if they are admitted, that their injuries will be recorded as being due to self-harm. If this chain of reasoning is correct, then the number of recorded self-harm hospitalisations in a particular time period is a valid (but not necessarily error-free) indicator of the number of actual acts of self-harm occurring in that time period. Nevertheless, in thinking about this process, measurement problems do become apparent: For example, not all acts of self-harm will result in a hospital admission (in fact, probably only a small minority do; Madge et al., 2008). Researchers using categorisations of behaviour need to attend to such problems: Just because a particular categorisation possesses some validity does not mean that the categorisation is error-free. A crucial issue is whether the sources of error in categorisations might be correlated with whatever explanatory variables are being studied. In the case of the self-harm hospitalisation data used in Study Three, there were several sources of likely errors in categorisation (e.g., acts of self-harm not being recorded as self-harm hospitalisations), as alluded to above, but little reason for concern that these variables might be correlated with temperature. As such, errors of measurement—or more specifically, of categorisation—may have not been a major problem bearing on the substantive findings in this particular instance. That will not always be the case.

6.2 Report Findings Using Unstandardised Effect Sizes

As noted in the methodological critique, another methodological strategy that can hamper the clear reporting of the magnitude of effects is the use of standardised effect sizes such as Cohen's *d*, correlation coefficients, standardised regression coefficients, and so forth. These standardised measures of effect size are intended to deal with two problems: measurement scales that have units of measurement with no clear and meaningful interpretation (Nakagawa & Cuthill, 2007; Wilkinson & Task Force on Statistical Inference, 1999), and the use of incomparable measurement scales across different studies (Baguley, 2009). Yet in reality, any given standardised effect size will typically be based on a metric (such as the sample standard deviation of the dependent variable) whose magnitude depends heavily on arbitrary methodological decisions by the researcher that are specific to the study at hand, and that harm the comparability of effect sizes across studies. Furthermore, as noted in the methodological critique, the standardised effect size calculated based on data from a sample is likely to be a biased estimate of the standardised effect size in the population, even if the corresponding unstandardised effect size is unbiased. The use of standardised effect sizes thus in no sense facilitates clear reporting of the magnitude of effects.

By instead using scales whose values have clear interpretations—even if this means simply classifying behaviour or individuals into clearly defined categories—psychologists can report unstandardised effect sizes that have clear and interpretable metrics. Such unstandardised effect sizes include mean differences and unstandardised regression coefficients. When working with a categorical dependent/response variable, these effect sizes may often be reported in multiplicative form: for example, as the percentage increase in the frequency of some behaviour given a particular change in conditions. Effect sizes that are communicated in this way are clearly interpretable. They can also be compared and synthesised across studies. For example, it would be entirely possible to perform a meta-analysis averaging estimates of the percentage increase in violent behaviour given a 1°C temperature increase. A meta-analysis does not necessarily require standardised effect sizes (see for example Dattilo & Kris-Etherton, 1992). On the other hand, if psychometric scales are used whose units of measurement are so difficult to interpret as to prompt concern about how the ensuing unstandardised effect sizes can be interpreted, this should be viewed as a prompt to consider alternatives to the measurement strategy used—*not* to switch to a standardised effect size, whose metric is likely to be biased and arbitrary.

6.3 Develop and Apply Theory that Facilitates Useful Prediction-Making

A recurrent theme in this thesis has been that psychologists do not have access to theory that is as useful for prediction-making as the body of theory applied by climatologists. The theory used by climatologists is capable of predicting (albeit with uncertainty) the magnitude and timing of future climatic changes. It can also be used to directly produce these predictions, without the use of statistical models trained on empirical data. It is unrealistic to expect that in the foreseeable future psychologists will produce quantitative theory of comparable sophistication and usefulness. Indeed, most psychological theory is not quantitative at all, but rather constituted of verbal statements about behaviour (Myung & Pitt, 2002). It is nevertheless the case that psychologists can develop and apply theory that *facilitates* the making of useful predictions—even if the theory applied is not capable of producing these predictions without help from statistical models trained on empirical data.

One situation in which a theory can be used to facilitate prediction is if the theory identifies causal effects that can be empirically estimated, and where knowledge of those causal effects can be used to generate useful predictions. For example, a theoretical framework describing how climate change may impact mental health was proposed by a group of Australian researchers (Berry et al., 2010). As discussed previously (see section 1.2.2.1), Berry and colleagues' framework suggests that adverse weather events occurring as part of climate change can impact mental health directly (e.g., by causing trauma); via economic, social and demographic impacts on communities that subsequently impact mental health; and via impacts on physical health that subsequently affect mental health. A strength of this model is that it specifies causal relationships between variables that can be estimated empirically in real-world contexts (albeit with necessary uncertainty) and that can be used to generate predictions about behaviours of genuine substantive interest.

As an example of how Berry and colleagues' (2010) model could be used to generate useful predictions, we could take a particular indicator of mental health that was studied in the empirical section of this thesis (suicide rates). In their model, one of the pathways via which climate change may impact mental health is via effects of climate change on economic productivity, which may subsequently affect mental health. The future impact of climate change on economic productivity can be estimated (see N. Stern, 2007), and so too can the impact of economic productivity on suicide rates (M. Berk, Dodd, & Henry, 2006; Chang, Gunnell, Sterne, Lu, & Cheng, 2009; B. Yang, Lester, & Yang, 1992). Obviously these estimates come with a non-trivial magnitude of uncertainty, and for present purposes it is not necessary to attempt a substantive synthesis of the research on these topics. Nevertheless, these estimates could be combined to make predictions about the effect of climate change on suicide rates—or at least of the effect occurring via the pathway of climate change effects on economic productivity. The theoretical framework can be applied in this way because the framework is made up of a sequence of proximal causal effects that can be estimated, and estimates of these proximal effects can be used to make predictions about the size of a distal effect (of climate change on suicide rates) that is of genuine substantive interest.

The advantage of drawing on theory for facilitating predictions is that it allows for the prediction of effects that may occur as a result of sustained climate change in the future, but that do *not* currently occur as a result of normal climatic variation. For example, normal temperature variation currently has little relationship with economic productivity (at least in wealthier countries; see Dell, Jones, & Olken, 2012). However, sustained climate change may have substantial effects on economic productivity (N. Stern, 2007). This means that the effect that sustained future climate change will exert on suicide rates via the pathway of impacting economic productivity is likely not captured by analyses of the effects of normal temperature variation on suicide. Provided that the effect of sustained climate change on economic productivity can be predicted (likely by drawing on work by researchers in other fields), and the effect of economic productivity on suicide rates can be estimated, then the size of the effect of climate change on suicide rates that may occur via the pathway of economic effects can actually be predicted. Such predictions are facilitated by the use of theory, and would not be possible simply by observing correlations between climatic variables and suicide rates. Use of theory in this way may also facilitate the study of impacts of climate change on psychological variables that occur slowly and that are mediated via wider societal processes and phenomena.

In sum, theories can be developed that facilitate the making of practical and useful predictions. However, to do so they need to be constructed in such a way that they specify causal effects that are estimable, and that can be used to make predictions about behaviours of substantive interest. To achieve this, psychologists may need to prioritise prediction (as opposed to explanation) when developing and applying theory. In general, the suggested focus on a theory’s capacity to make *useful* predictions aligns well with an instrumentalist stance toward psychological theory (see Cacioppo, Semin, & Berntson, 2004).

6.4 Take Into Account the Role of Time

Climate change is a problem that has evolved and will continue to evolve over a long time frame, with causal antecedents and consequences often separated by substantial periods of time. Predictions about the effects of climate change on psychological variables will be of most value if they can specify when these effects occur. Such predictions require the use of theory and analyses that explicitly take into account the role of time. Psychologists studying climate change impacts will need to be aware that the effects of exposure to particular environmental conditions may differ substantially depending on the time of exposure. For example, sustained changes to mean temperatures may affect human behaviour in different ways than do random short-term changes in temperature. This was a point that was made in the empirical articles presented in this thesis.

One important way in which psychologists may study the effects of different timeframes of exposure to particular environmental conditions (e.g., warmer temperatures) is by studying the relationships between human behaviour and different components of variation in environmental conditions. The idea that series of observations can be separated into different components of variation is widely recognised in fields such as econometrics that make extensive use of time series analysis (e.g., P. S. Mann, 1994), but is less widely appreciated in psychology. The analysis of the effects of different components of variation in temperature was extensively used in the empirical articles in this thesis. Doing so facilitated an examination of whether brief exposure to warmer temperatures (as in the analyses of irregular daily temperature variation), had similar effects to more sustained exposure to warmer temperatures (as in the analyses of seasonal and geographical variation in temperature). Finding ways to study the effects of more sustained exposure to particular environmental conditions is important for psychologists interested in studying the impacts of climate change given the limited timeframe of observations that may be available. In discussing the study of the effects of different components of variation in climatic conditions, two components of variation deserve particular attention: geographical variation, and long-term trends.

6.4.1 Studying geographical variation in environment and behaviour.

By studying the differences in behaviour between populations living in different geographical areas, psychologists can make tentative observations about the effects of *sustained*, long-term, differences in environmental conditions. There are, however, two important challenges that arise when studying the effects of geographical variation in environmental conditions: confounding, and small sample size.

The first problem with studying the effects of geographical variation in environmental conditions is the vulnerability of such analyses to confounding. Human populations living in different areas will differ in terms of many social, economic, demographic and cultural variables that may impact the behaviours being studied (e.g., assaults, suicide, etc.). Many of these extraneous variables may have some statistical relationship with the environmental variables being studied as causal inputs. This is a difficult problem to rectify firstly because there are many such extraneous variables, making it difficult to identify, measure and statistically control for the most important potential confounds. Secondly, determining whether a given extraneous variable is truly a confounding variable requires knowledge about the causal effects of variables on one another rather than just the statistical relationships between them. For example, Hsiang et al. (2013) argue *against* controlling for socioeconomic variables when examining the effects of temperature on conflict, because it is difficult to exclude the possibility that temperature itself affects such socioeconomic variables—in which case they would not have the formal characteristics of confounding variables[25]. Statistically controlling for a variable that is correlated with the independent variable but also affected *by* the independent variable (thus not having the characteristics of a confounding variable) can bias the estimate of the effect of the independent variable (McNamee, 2003; Weinberg, 1993). Researchers attempting to analyse the effects of geographical variation in environmental conditions (e.g., temperature) on psychological variables will therefore have to put a great deal of thought into which statistical controls to apply. In many cases this will mean not just examining correlations between potential confounding variables and the independent and dependent variables studied, but also examining the related literature to determine the likely causal mechanism of relationships found: That is,

---

[25] A confounding variable is one that is correlated with the independent variable, not affected *by* the independent variable, and that affects the dependent variable (see McNamee, 2003; Weinberg, 1993). In contrast, an extraneous variable is any variable other than the independent variable that affects the dependent variable.

based on the current state of knowledge about causal effects in the domain studied, which variables are most likely to have the characteristics of confounds?

The second challenge when examining the effects of geographical variation in environmental conditions is that of dealing with small sample sizes and corresponding large quantities of sampling error. Although the methods used may differ somewhat from study to study, in general a recording of behaviour within one geographical area (no matter how long the time period) represents a single observation in an analysis of the effects of geographical variation in environmental conditions on human behaviour. This meant, for example, that the studies of geographical variation in temperature and the incidence of violent behaviour presented in this thesis had sample sizes of just 66 or 67 districts (this being the number of districts in New Zealand, depending on whether or not the Chatham Islands were included). The resulting inferences were thus subject to a great deal more uncertainty than those produced in the analyses of irregular variation in temperature, in which the sample sizes involved were thousands of dates. Specifically, the confidence intervals for the effect of geographical variation in temperature on the three types of violent behaviour studied were very wide. In this case the sample of geographic regions studied (New Zealand) was not ideal, covering a relatively small geographical area itself divided into a reasonably small number of regions. The obvious solution to this problem would be to sample data from a wider geographical area, thus reducing sampling error in geographical analyses. Sampling from a wide geographical area also has the added benefit of allowing researchers to study the effects of a wider range of values in the environmental predictors of interest (e.g., a wider range of temperatures), thus avoiding any problems with restriction of range. It also makes it more likely that populations in areas particularly vulnerable to climate change (e.g., those that are economically, socially and politically marginalised; see IPCC, 2014b) will be studied. In comparing populations across large and diverse geographical regions, however, researchers will need to be particularly wary of the problem of confounding by social, economic, demographic and cultural variables (as discussed above). They will also need to attend carefully to practical problems such as differing data collection practices across different geographical areas (a problem that came up in Study Three; see section 4.3.2).

6.4.2 Studying trends in environment and behaviour.

One other component of variation in environmental conditions also deserves some special consideration with respect to its relationship with human behaviour. This component of variation is that of long-term *trends* in environmental conditions. On the face of it, an ideal way to study the likely effects of sustained changes in climate on human behaviour would be to obtain long series of observations of environmental conditions and the behavioural variables of interest. One could then remove irregular short-term variation in each series (e.g., by aggregating data over years or decades), and then study how the long-term trends in these environmental and behavioural variable relate to one another. Such studies of long-term trends have occasionally been conducted in the temperature-violence literature (e.g., Helama et al., 2013; Holopainen et al., 2013).

Unfortunately, this approach comes with several challenges. The first is simply the practical difficulty of obtaining long and consistently recorded series of observations of the variables studied. This task may be feasible for meteorological variables, but is often difficult for behavioural variables. The second challenge is the problem that the climatic stability of the Holocene epoch (Dansgaard et al., 1993) means that there may be relatively little variation in long term trends in meteorological variables to study when analysing data from the past. This is despite the fact that there may be a great deal of fluctuation in meteorological variables over short periods of time. In effect, this is a problem of restriction of range that impacts the capacity to discover relationships between climatic and behavioural variables.

Finally, when comparing human behaviour across long periods of time, there is again the possibility of confounding due to changing social, demographic, economic and cultural conditions—much as is the case for analyses of geographical variation. These problems of confounding can largely be circumvented when analysing the effects of short-term (e.g., daily) variations in environmental conditions by the simple expedient of controlling for the effect of time period. For example, in many of the empirical analyses reported in this thesis, controls were applied for the year of observation, providing a flexible control for any potential confounds that varied reasonably slowly over time. But when analysing long-term trends, it is not so simple to control for time-trending confounds in this simple fashion. If using series of annual data, for example, one may not be able to control for the year of observation while retaining an over-identified model (depending somewhat on the method used to apply the control). This means that when analysing long-term trends, just as when analysing geographical variation in environmental conditions and behaviour, researchers will need to specifically measure and control for potential confounding variables. Of course, obtaining observations of potential confounding variables over long periods of time may itself present difficulties. The study of long-term trends in environmental and behaviour variables thus could have some benefits, but comes with some significant challenges.

6.5 Improve the Modelling and Reporting of Uncertainty

Conclusions about the impacts of climate change on human behaviour will be most useful when they are accompanied by accurate information about the uncertainty surrounding those conclusions. As mentioned in the methodological critique in section 5.2.3, psychologists tend to estimate and report only the uncertainty arising from a very limited range of sources (principally sampling error). Yet uncertainty with respect to conclusions and predictions can arise from many sources. A number of such sources are identified in section 5.2.3. The effects of some sources of uncertainty may be modelled and reported quantitatively. The effects of other sources of uncertainty may be harder to quantify. In such cases, the best researchers may be able to do is simply to verbally acknowledge the presence of unquantified sources of uncertainty.

In contrast to the sequence of ideas presented in the methodological critique, here I will discuss firstly the topic of how uncertainty is reported, and then only later move on to discussing how the effects of different sources of uncertainty should be managed. By changing how uncertainty is reported in psychological studies, it may be possible to account for a wider range of sources of uncertainty.

6.5.1 Report uncertainty using interval estimates.

In section 5.2.1 I argued that it is more important for psychologists to report estimates of the *size* of effects of climate change than to simply identify their direction. An implication of this point is that estimates of uncertainty may most usefully be reported in the form of intervals within which we can say that a true effect lies with a particular degree of certainty. Such interval estimates are clearly more useful than *p* values. *p* values are a form of information about uncertainty with no clear intuitive interpretation, and that do not directly answer any question likely to be of real substantive interest to a user of climate change research. The argument for using interval estimates of uncertainty—such as confidence intervals and credible intervals—has been made repeatedly elsewhere (e.g., Cumming, 2013; Gardner & Altman, 1986; Gill, 1999), and I will not linger on it here. A question that is not as readily resolved, however, is whether interval estimates should be reported in the form of Bayesian credible intervals, frequentist confidence intervals, or some other form of interval estimate. A discussion of the merits of several approaches follows.

6.5.1.1 Bayesian analysis and credible intervals.

As discussed in the methodological critique, a problem with confidence intervals is that a single specific confidence interval has no clear intuitive interpretation—or at least not if interpreted strictly in frequentist terms. As mentioned in section 5.2.3.4.2, we can say that if we conducted a very large number of repeated independent studies estimating some parameter, and calculated a 95% confidence interval each time, then approximately 95% of these intervals will include the true parameter. But from a frequentist perspective we cannot say that there is a 95% probability that an actual individual interval includes the true parameter (Hoekstra et al., 2014).

An alternative interpretation of probability is the more flexible Bayesian interpretation, in which probability refers to a quantity of uncertainty or degree of belief (Wagenmakers, 2007). The Bayesian interpretation of probability allows for useful probability statements to be made about many kinds of statements, including statements about whether particular hypotheses are true or false, or whether a parameter actually falls within a given interval. The cornerstone of Bayesian data analysis is Bayes theorem (Bayes & Price, 1763). Bayes theorem shows how we can take a *prior* probability distribution (which describes our beliefs or knowledge before seeing the data collected) and then update the prior distribution using the data collected. The result is a *posterior* probability distribution. The posterior probability distribution can directly indicate the probability that a particular hypothesis is true (for a binary proposition), or which values of a parameter are most probable (for a continuous parameter).

Unlike frequentist analysis, Bayesian analysis permits the calculation of interval estimates for which one can make a statement such as "there is a 95% probability that the parameter lies within the interval" (Rindskopf, 2013, p. 323). These intervals are known as credible intervals. Credible intervals can be calculated in a variety of different ways (Bååth, 2014), with the highest density interval being a particularly important example (see Kruschke, 2010a). Despite the clear and intuitive interpretation of Bayesian credible intervals, Bayesian analyses have been extremely rare in the temperature violence area. A rare example is an Australian study that reported a credible interval of 0.73–3.82% for the effect of a 1°C temperature increase on suicide rates (Qi, Hu, Mengersen, & Tong, 2014). Bayesian analyses are also reported in the supplementary materials section of Study Three in this thesis (see section 11.3.4).

6.5.1.2 Confidence intervals with Bayesian interpretations.

Despite the apparent advantages of Bayesian analysis for communicating research results, one possible justification for instead reporting frequentist confidence intervals is that, given weak prior information, frequentist confidence intervals closely approximate Bayesian credible intervals[26] (Greenland & Poole, 2013). In fact, Greenland and Poole note that a credible interval and the corresponding confidence interval will coincide *exactly* when the prior probability distribution places equal odds on all values of the parameter (e.g., a continuous uniform prior for a regression slope with limits of $-\infty$ and ∞). Such a prior would indicate a state of complete ignorance before seeing the data collected, and might also be described as a “non-informative” prior. Because frequentist confidence intervals are based only on the data at hand, the more information provided by the sample data (e.g., the larger the sample size), and the less informative the prior, the more closely a frequentist confidence interval will approximate the associated Bayesian credible interval. In the Bayesian analyses of the relationship between self-harm and temperature reported in this thesis, the credible intervals calculated were very similar to the corresponding confidence intervals calculated via frequentist analysis (see Table 21 in section 11.3.4). This was even despite the use of reasonably informative priors in the Bayesian analysis (specifically, the priors used indicated an assumption that smaller effects were more plausible than larger ones).

If frequentist confidence intervals approximate Bayesian credible intervals quite well, researchers might be justified in choosing to calculate frequentist confidence intervals on pragmatic grounds. One of these pragmatic grounds is the fact that Bayesian analyses generally rely on Markov Chain Monte Carlo estimation (see Kruschke, 2010b), meaning that estimation is computationally intensive and can take a

[26] Greenland and Poole (2013) use the term “posterior probability interval” (p. 62) instead of “credible interval”.

significant quantity of time. Another pragmatic reason for using frequentist analysis is that Bayesian analysis is more demanding of the data analyst, requiring more programming skill to estimate models than is the case for common frequentist analyses that can easily be accomplished using programs with graphical user interfaces (e.g., SPSS). Even where the analyst is comfortable with a Bayesian approach, the increasingly short word limits imposed by some journals may make it difficult to find space to introduce new statistical concepts, thus providing an incentive to use familiar analyses with minimal explication.

By using frequentist analyses, researchers may also sidestep the difficult task of the specification of prior probability distributions. Indeed, the notion that the results of an analysis may depend at least somewhat on the prior subjectively selected by the researcher is one reason why some researchers are uncomfortable with the Bayesian approach (see the discussions by Efron, 1986, and Gelman, 2008). Whether this discomfort with the apparent subjectivity of a Bayesian approach is actually a legitimate objection is somewhat dubious: Even a frequentist analysis involves a great number of decisions made subjectively by the researcher (Gill, 2015). Indeed, this navigation of "the garden of forking paths" has problematic implications for the interpretation of frequentist statistics (Gelman & Loken, 2013, p. 10). Bayesian analyses do depend on subjective information specified in the prior probability distribution, but they do so in a transparent fashion, and in such a way that the analyses can easily be replicated with different priors (Kruschke, 2010a). But from a pragmatic perspective, avoiding having to justify the use of subjective priors to editors, reviewers and readers unfamiliar with Bayesian analysis may appeal to researchers, particularly when limited space for communication is available.

Some researchers might thus choose to use frequentist analyses and confidence intervals, but interpret their confidence intervals in a Bayesian fashion. This approach combines the pragmatic advantages of frequentist analysis (for the researcher), with the direct and intuitive interpretability of Bayesian statistics (for the reader). However, this approach only works insofar as researchers are actually prepared to *communicate* how and why the confidence intervals that they report can be interpreted in an intuitive (Bayesian) fashion. For example, consider the confidence interval of 0.012–0.017 for the effect of irregular variation in temperature on assault incidence reported in this thesis (see section 2.4.1.3). This interval could be interpreted as such:

> Assuming that all effect sizes could be considered equally probable prior to observing the data at hand, and assuming a lack of any uncontrolled confounders or breaches of distributional assumptions, there is a 95% probability that the true size of the effect of a 1°C increase in temperature in New Zealand is an increase in assaults of between 1.2% and 1.7% (based on the Bayesian interpretation of confidence intervals suggested by Greenland & Poole, 2013).

This interpretation shows how the confidence interval can be interpreted in a Bayesian fashion, as discussed above. It also explicitly highlights the dependence of the statement made on a variety of assumptions. Unfortunately, researchers reporting confidence intervals rarely provide explicit interpretations of the intervals reported (Cumming et al., 2007). If interval estimates (whether Bayesian or frequentist) are reported without clear information about how these estimates can be interpreted, it will be difficult for readers and stakeholders to understand the information about uncertainty provided.

6.5.1.3 Consider reporting interval estimates not solely based on probability distributions.

The two subsections above discuss the relative merits of two kinds of interval estimates that rely on different interpretations of the term "probability". However, both confidence intervals and credible intervals are still based on probability distributions. Indeed, quantitative statements about uncertainty in psychology are almost always based directly on probability distributions. This is the case whether the probability distribution utilised is a sampling distribution (in frequentist analysis) or a posterior probability distribution (in Bayesian analysis). In some cases, however, it may be desirable to produce an interval estimate conveying a magnitude of uncertainty even when it is not feasible to base this statement solely on a single probability distribution. I will explain below why this approach may be helpful to psychologists studying climate change impacts.

The approach of reporting interval estimates of uncertainty that are not solely based on probability distributions is sometimes used in climatology. In the IPCC's fifth assessment report, temperature projections are provided for a range of timeframes and emissions scenarios (see IPCC, 2013b). These projections are provided in the form of both point and interval estimates. For example, given the representative greenhouse gas concentration scenario 8.5 (RCP8.5; a scenario with increasing greenhouse gas emissions throughout the 21st century), the point estimate of the global temperature increase over the 21st century is 3.7°C, with a *likely* range of 2.6°C to 4.8°C. The limits of 2.6°C to 4.8°C are the 5th and 95th percentiles of the estimated temperature increase across 39 climate models[27]. In this sense the interval is similar to a 90% confidence interval. However, rather than being described as a 90% confidence interval, the interval is described as a *likely* range, where *likely* is formally defined as meaning a probability of between 66 and 100%. In other words, the projection should be read as indicating that we can be between 66 and 100% certain that if future emissions take the trajectory described in the RCP8.5 scenario, then the temperature increase over the 21st century will be between 2.6 and 4.8°C.

Importantly, this statement about uncertainty is not based purely on a probability distribution. The estimates of temperature increase produced by the climate models have a probability distribution, and this distribution was used to produce the interval reported. However, the statement made about uncertainty is not based exclusively on this probability distribution (if it was, we might indeed refer to the interval as a 90% confidence interval). Rather, the statement made is also based on the expert judgment of the IPCC authors and their knowledge about sources of uncertainty that are *not* captured in the climate models used. Despite not being solely based on a probability distribution, the statements made are nevertheless quantitative: The prediction made is quantitative (a temperature increase of between 2.6 and 4.8°C), and so too is the statement about the magnitude of uncertainty present (the probability of the true temperature increase falling within this interval is between 66 and 100%).

Obviously, providing quantitative statements about uncertainty that are not based directly on probability distributions has its drawbacks. A cost of this approach is that the statements made about uncertainty are less precise. In the above example, the probability that the true temperature increase given the RCP8.5 scenario will be between 2.6 and 4.8°C is not assigned a single value (such as 90%); rather, the probability is believed to fall within the broad range of 0.66 to 1. Another problem is that judgments about uncertainty become more subjective, being based on the judgments of experts. However, the benefit of making quantitative statements about uncertainty that are not based directly on probability distributions is that doing so

---

[27] Albeit that the percentiles were calculated from the mean and standard deviation of the estimates along with an assumption of normality, rather than directly by ranking the estimates.

permits an accounting for sources of uncertainty that cannot be easily captured within a single probability distribution.

Furthermore, the possibility of using an interval estimate that is not based on a single probability distribution opens up the possibility of an important strategy. This strategy is that of estimating the size of effects or relationships across multiple models making different assumptions and/or using different data, and reporting an overall interval estimate based on variability in results both within and across models. For example, in Study Two it would have been possible to report an overall interval estimate for the effect of temperature on suicide by finding the lowest 95% confidence interval limit across the models estimating this effect in the study, as well as the highest 95% confidence interval limit, and using this overall interval as an indicator of uncertainty. This overall interval estimate would not be based directly on a probability distribution and thus could not be interpreted directly as a 95% confidence interval. It would nevertheless represent an interval estimate that captures a much greater range of sources of uncertainty than an interval estimate from a single model. Indeed, several sources of uncertainty with respect to statements about climate change impacts are potentially quantifiable, but difficult to incorporate within a single probability distribution or model. These include some forms of uncertainty about distributional assumptions in statistical modelling (e.g., when there is more than one distribution that might plausibly be used for the error term); uncertainty about the form of relationships (i.e., where several functional forms are plausible); and uncertainty about the values that causal inputs will take in future.

When reporting an interval based on a range of models rather than a single probability distribution, the magnitude of the uncertainty about whether the across-model interval estimate actually includes the true impact of interest could then be indicated as a range of values. This could even be accomplished by using the uncertainty terms and definitions used by the IPCC (e.g., 66–100% = likely; 90–100% = very likely; 99–100% = virtually certain; see Mastrandrea et al., 2011). This would require researchers to make necessarily subjective judgments about the importance of any remaining sources of uncertainty.

To some limited extent the approach of estimating uncertainty across multiple models was taken in the discussion section of Study One, where an interval estimate of the size of the effect of a 2.1°C increase in temperature on assaults was provided showing how the estimated size of this effect varied across a number of models estimated (see section 2.5.2). This said, the range of estimates discussed was the range of point estimates across models, not the range of confidence interval limits. It was also not accompanied by a statement indicating how certain we could be that the true effect fell within the interval provided.

6.5.2 Take into account more sources of uncertainty.

The preceding discussion of the ways in which uncertainty can be reported opens the door for a more specific discussion of how individual sources of uncertainty with respect to conclusions about climate change impacts might better be accounted for in psychological research. These sources of uncertainty are now discussed in turn.

6.5.2.1 Uncertainty about the validity of distributional assumptions.

As noted in the methodological critique, inferential data analyses are inevitably based on assumptions, and uncertainty with respect to the validity of these assumptions is not usually taken into account in model outputs (such as confidence intervals). While distributional assumptions are generally not required simply in order to describe variables and relationships observed within a sample, they are required whenever a researcher wishes to use a statistical model to make inferences from a sample to a population. They are likewise required when a researcher wishes to make inferences from a sample of observations to the causal effects that gave rise to those observations. In the latter case, additional assumptions are also necessary, as discussed previously. Researchers can obviously investigate the validity of distributional assumptions using statistical or graphical methods (see for example J. Cohen, Cohen, West, & Aiken, 2003; and Gelman & Hill, 2007). However, such investigations are very unlikely to result in complete confidence that the assumptions of a particular method are actually met. Uncertainty with respect to the validity of distributional assumptions will thus add to the uncertainty surrounding substantive conclusions.

One way to deal with uncertainty arising due to distributional assumption breaches is obviously to reduce the reliance on such assumptions, such that whether or not a particular[28] assumption holds true is no longer a source of uncertainty. For example, while a conventional application of the Poisson model for count data assumes that the variance of the response variable is equal to its predicted value for a given

[28] It is effectively impossible to avoid making *any* distributional assumptions as part of a statistical model, but specific assumptions may be dispensed with in some cases.

combination of values of the predictors, this assumption was avoided in the analyses of the relationship between temperature and assault[29] reported in this thesis. This was accomplished in most of the models applied by incorporating a freely estimated error term. This error term allowed for the presence of overdispersion (i.e., variance in excess of the mean). More generally, some distributional assumptions commonly relied on in statistical analyses in the past can be sidestepped in contemporary practice due to the possibility of bootstrapping and permutation tests (see Hesterberg, Moore, Monaghan, Clipson, & Epstein, 2002). These tests can be used to empirically approximate the sampling distribution of coefficients without some of the distributional assumptions (e.g., that of normally distributed errors) that were required to produce computationally tractable analyses in the past.

In other cases, it may be impossible to completely avoid a problematic distributional assumption, but it may be possible to estimate different models that make different distributional assumptions. An interval estimate for an effect can then be calculated using the results of multiple models making different assumptions. This approach uses the general strategy of using an interval estimate that is not based on a single probability distribution. For example, analyses of the effects of geographical variation in temperature on self-harm reported in this thesis (see section 11.3.1) were reported using models making different assumptions about the relation between the conditional mean and conditional variance of self-harm incidents. The two models compared were the negative binomial model and the quasi-Poisson model. Simply reporting these different approaches gave some indication of the uncertainty arising due to uncertainty about which distributional assumption was more reasonable. This communication of uncertainty could have been improved even further by reporting an interval estimate for the effect of temperature calculated across the results from both models.

6.5.2.1.1 Uncertainty due to the sampling method.

As noted in the methodological critique (section 5.2.3.1.1), the methods used to quantify uncertainty in psychological research typically implicitly assume that simple random sampling has been used. This assumption rarely holds in practice: Even when

[29] This assumption appeared to be much more plausible in the suicide and self-harm studies (Studies Two and Three), where there was little evidence of dispersion in excess of that predicted by the Poisson model. As such the simpler Poisson model was used for most of the analyses in these two studies.

random sampling from a well-defined population is attempted, the lack of a perfect response rate will inevitably mean that the actual sample of data obtained is not a genuine random sample. Low response rates are a ubiquitous problem for researchers attempting to obtain probability samples (see Singer, 2006). It is in fact possible to use statistical methods (most notably multilevel regression and poststratification) to increase the accuracy of estimates obtained from a non-probability sample (see Gelman & Carlin, 2001; W. Wang, Rothschild, Goel, & Gelman, in press). Conceptually, this is accomplished by adjusting the statistical estimates obtained for known differences between the sample and the population (e.g., in terms of demographic factors). While this method is becoming of increasing importance in political science (Buttice & Highton, 2013), it is rarely used in psychology. Psychologists wishing to use sample data to draw inferences about the effects of climate change on a broader population could consider using multilevel regression and poststratification to improve the representativeness of their estimates. In cases where uncertainty is quantified based simply on an assumption of random sampling, researchers should qualitatively indicate that additional uncertainty arising due to the sampling method used exists and has not been quantified.

6.5.2.1.2 Uncertainty arising due to problems with measurement.

One particularly important source of potential distributional assumption breaches is the presence of measurement error. Uncertainty surrounding conclusions that arises due to uncertainty about the reliability and validity of measurements can be difficult to account for quantitatively. Yet measurement error can have a profound effect on estimates of effects and relationships. As noted previously in section 5.2.3.1.2, common statistical methods such as the linear regression model take into account the possibility of random (unsystematic) measurement errors in the response variable, but assume that this measurement error is uncorrelated with the predictors (Williams, Grajales, & Kurkiewicz, 2013). They furthermore assume that the predictors themselves are measured entirely without error (Montgomery et al., 2001). Rarely will there be any strong case to believe that these assumptions are fulfilled.

When there is random (unsystematic) measurement error in the predictor variables—error that is uncorrelated with the response variable observations—measurement error can be corrected for by simple formulae (Schmidt & Hunter, 1999; Spearman, 1904), or more generally via the use of errors-in-variables models

(Wansbeek & Meijer, 2003). When correlated measurement error is present, the situation becomes more complicated (and more unpredictable; see Williams, Grajales, & Kurkiewicz, 2013).

Correlated measurement error can be quantified and accounted for by using structural equation modelling (Bollen, 1989; Kline, 2005). This may be a useful strategy, but works only to the extent that the researcher is able to specify an overidentified[30] model that accurately depicts the relationships between variables and measurement error terms. Given a large enough sample size, any overidentified structural equation model can be shown to be an inaccurate approximation of reality according to a strict criterion such as the chi-square test (McDonald & Marsh, 1990). If the structural equation model applied is incorrect, then there will be no assurance that the parameter estimates it provides correctly account for measurement error. In most cases, even if complex techniques for dealing with measurement error such as structural equation modelling are applied, there will still be some unquantifiable uncertainty about substantive claims arising due to uncertainty about measurement error. This source of uncertainty should at least be acknowledged when describing a study (even if it cannot be quantitatively modelled).

6.5.2.1.3 Uncertainty about the form of relationships.

When specifying a statistical model there is usually some uncertainty about the forms of the relationships specified (linear, non-linear, etc.), as mentioned in the methodological critique (section 5.2.3.1.3). Uncertainty about the form of relationships can be a source of quantitative uncertainty with respect to the size of causal effects estimated from observed data. It can similarly be a source of uncertainty with respect to the magnitude of changes to behavioural variables that may arise due to climate change in the future. The specification of an incorrect form for the relationship between variables can also result in breaches of distributional assumptions in statistical models. While there are many true forms that any given relationship between two or more quantitative variables might take, this source of uncertainty can at least be partially addressed by two strategies.

One strategy is simply to expend some effort on using theory and/or empirical observations to determine which relationship forms are most plausible in a given

[30] An over-identified model is one that is less complex than the data it is intended to explain. See Kline (2005) for a more detailed discussion.

setting. For example, in the empirical studies reported in this thesis, strategies such as non-parametric regression using loess smoothing and spline models were used to investigate the form of relationships between variables. In general, a linear form seemed to closely approximate the relationships estimated in this specific context.

Another strategy is to generate claims about the size of the effect of a particular climatic change based on models making a *range* of plausible assumptions about the form of the relationships of interest (e.g., about the relationship between temperature and violence). For example, as discussed earlier, studies investigating the relationship between temperatures and assaults have claimed evidence both for a linear form of the relationship (Anderson, 1989; Bushman et al., 2005), as well as for a non-linear and non-monotonic form (Cohn & Rotton, 1997, 2005; Gamble & Hess, 2012). It would be possible to use a given dataset to estimate models taking both linear and non-linear (e.g., polynomial or spline) forms and to generate predictions about the effect of a given temperature change using both models. This is another case where reporting an interval estimate based on the results of multiple models (and not a single probability distribution) could be a valuable way to report the magnitude of uncertainty arising due to a wider range of sources.

6.5.2.2 Uncertainty about the validity of the assumptions necessary to make causal inferences.

When making causal inferences (e.g., about the effect of climate change on some behavioural variable) using empirical data, the possibility of breaches of the assumptions necessary for such causal inferences is an important source of uncertainty. These two assumptions are the stable unit treatment value assumption (Rubin, 2005) and the strong ignorability assumption (Rosenbaum & Rubin, 1983). The most obvious potential problem is that of an omitted confounding variable causing a breach of the strong ignorability assumption. This is an especial concern when making causal inferences from sources of data other than true experiments, as may often be the case when studying the impacts of climate change. In theory it is possible to adjust estimates of causal effects for the possibility of confounding using a sensitivity analysis (McCandless, Gustafson, & Levy, 2007). However, such adjustments can be complex to produce and report, and to large extent rely on the investigator's ability to subjectively determine what magnitude of bias due to unmeasured confounding is plausible. To my knowledge, a sensitivity analysis of this kind has never been used in a study concerned with the psychological impacts of climate change. In general, researchers may be restricted to qualitatively acknowledging the presence of this source of uncertainty. Such an acknowledgement should preferably be accompanied by a discussion of how likely the possibility of an omitted confounding variable or a breach of the stable unit treatment value assumption is. The discussion sections of the three empirical studies reported in this thesis, for example, gave especial attention to the issue of confounding.

6.5.2.2.1 Uncertainty due to varying effects.

As noted in the methodological critique, if the effect of a particular "treatment" (such as exposure to a particular temperature) varies across units, this has important and problematic implications. Varying treatment effects can result in additional uncertainty about the *average* effect of a particular change in conditions. It can also result in uncertainty about the size of the effect occurring in a particular time, place, or population. There are several potential solutions to these problems, but none of them are trivial to implement. J. Kim and Seltzer (2011) suggest a method for identifying varying response to treatment based on heterogeneity in the residual variance across groups in a mixed/multilevel model. Aronow and Samii (in press) show how an estimated average treatment effect can be adjusted such that all units are weighted equally, improving the representativeness of this estimate. Another solution is to model varying treatment effects as part of a mixed model (Gelman, 2004; Gelman & Huang, 2008). The importance of this problem and the appropriate response may depend on whether the specific research problem calls for an estimate of an average effect or an estimate of how an effect varies across people, times or places. Similarly, the appropriate response may also depend on the degree to which the effect estimated seems likely to vary across the specific units studied. Nevertheless, psychological researchers studying the impacts of climate change should at least be aware of the fact that the effects of climate change may vary across people, times, and places, and consider modelling how effects vary. For example, in the assaults article reported in this thesis, an average effect of temperature on recorded assaults was estimated across Auckland, Wellington, and Canterbury, but so too was the effect *within* each individual region (see section 2.4.1.3).

6.5.2.3 Uncertainty about whether model parameters will change over time.

As noted in the methodological critique, one source of uncertainty relating especially to predictions about the future is uncertainty about whether the parameters of the model used to generate predictions might change over time (e.g., if the effect of temperature on violence will change in the future). In cases when observations of the environmental and behavioural variables of interest are available for a reasonably long period of time, it may be possible to empirically investigate the degree to which the parameters have changed in the past (e.g., is the effect of temperature on violence invariant across years?) Of course, such investigations will only provide a tentative idea of how much model parameters (e.g., causal relationships) may change in the future. It may therefore be important to qualitatively communicate the presence of this source of uncertainty when reporting predictions about the future, with an acknowledgement that the effects of this source of uncertainty are difficult to quantify.

6.5.2.4 Uncertainty about causal inputs.

Another source of uncertainty with respect to predictions is uncertainty about causal inputs. For example, psychologists may wish to make predictions about human behaviour in a future, warmer, world. However, if these predictions are based on statistical models estimating the causal effects of climatic variables, the predictions made will necessarily be contingent on the magnitude of actual future changes to these climatic variables.

This is another source of uncertainty that might be usefully dealt with by reporting interval estimates of uncertainty that are not based on a single probability distribution. Psychologists could produce estimates of the effect of climate change on a behavioural variable based on different plausible values for the input variables (e.g., temperature change). They could then calculate an interval estimate for the predicted behavioural variable for each value of the input variable, and then report the overall range of the interval estimates. In this way, uncertainty arising due to uncertainty about the future trajectory of the causal input variable will be incorporated in the resulting interval estimate.

In doing so, it would be important to select a range of values for the causal inputs (e.g., climatic variables) that represents a plausible range of scenarios given current knowledge about the future. For example, predictions about global temperature change based on a range of different models and assumptions about future emissions can be obtained from IPCC reports (e.g., IPCC, 2013a), and then used as a basis to produce predictions about impacts on behavioural variables.

Using an estimated model to produce predictions under a range of plausible future climate scenarios is especially important when the model takes a complex form.

The models used to predict global warming impacts on violent crime rates in a group of well-known studies by the Anderson group (Anderson, 2001; Anderson et al., 2000; Anderson & DeLisi, 2011) were simple linear models of violent crime rate regressed on temperature and other control variables. Converting a hypothesised temperature increase into a predicted impact on the violent crime rate could be accomplished for these models as a simple mental exercise for a reader (i.e., simply by multiplying the estimated effect of temperature by the hypothetical temperature change). This translation step would similarly be fairly simple for the models reported in this thesis. But for models that include complexities such as the non-linear effects reported by Gamble and Hess (2012), the translation from model to prediction is obviously more difficult, even for the researchers estimating these models. For example, Gamble and Hess concluded that their finding that the relationship between temperature and acts of interpersonal violence such as aggravated assaults became negative at temperatures above 32.2°C implied that higher temperatures due to global warming "are not likely to be accompanied by markedly higher rates of violent crime and associated increased incidence of injury and death" (p. 239). In reality, data simulated according to their model actually implies that higher temperatures would produce higher average assault rates for all temperature increases that are remotely plausible over the next century (Williams, Hill, & Spicer, 2013; see Appendix D). As such, taking care to calculate and report what an estimated model actually implies about the future under a range of assumed values for the causal inputs specified in the model is an important task.

6.5.2.5 Uncertainty due to stochastic behaviour of the predicted variables.

A final source of uncertainty with respect to predictions mentioned in the methodological critique was that of the stochastic (probabilistic) behaviour of the variables likely to be predicted. This source of uncertainty needs to be considered separately only when attempting predictions about the actual values of variables at some specific time in the future. It is not relevant when simply attempting to make predictions about the *effect* that a particular climatic change might exert. Thus far, psychologists studying climate change impacts have generally only attempted to make the latter type of prediction (e.g., Anderson et al., 2000; Anderson & DeLisi, 2011). This was also the case in the empirical studies reported in this thesis.

Whether psychologists are able to produce predictions of the former variety—about actual behaviour occurring in the future—with any useful degree of accuracy in the climate change context is an open question. However, if such predictions are attempted it will be important to account for uncertainty occurring due to stochastic variability. The study of climate change impacts is an area of psychological inquiry in which accurate predictions about behaviour in the future might be especially useful (for the purpose of guiding adaptation efforts). If a prediction of the actual value of some behavioural variable at some time in the future is attempted, uncertainty occurring due to the stochastic nature of the predicted variable can be dealt with by using prediction intervals (see Chatfield, 1993). These prediction intervals may take either frequentist or Bayesian form. While confidence and credible intervals reflect uncertainty only with respect to the values of model parameters, prediction intervals indicate the magnitude of uncertainty about the values of actual future observations. Prediction intervals achieve this by taking into account both uncertainty about the value of model parameters as well as additional uncertainty with respect to the values of future observations occurring due to stochastic variability.

7 Conclusions

This thesis comprised both an engagement in empirical research concerned with the psychological impacts of climate change as well as a critique of the suitability of psychologists' mainstream methodological strategies for engaging in this research area. In the empirical component, I found that irregular daily variation in temperature had a modest positive effect on the incidence of three forms of violence in New Zealand: assaults, suicides, and acts of self-harm resulting in hospitalisation. However, the effects of more sustained exposure to warmer temperatures (e.g., the effects of seasonal and geographical variation in temperature) were less consistent and subject to much more uncertainty. It was also impossible to rule out the possibility that climate change may affect the incidence of intra- and interpersonal violence via mechanisms other than those that drive current short-term relationships between temperature and violence. It was therefore difficult to draw firm conclusions about how climate change will affect human violence in the future—or at least not on the basis of conventional psychological research methods.

In the methodological critique, I went on to discuss several core problems with psychologists' favoured quantitative research strategies that limit our capability to produce effective and useful research concerned with the impacts of climate change on human behaviour. These problems include measurement and analysis strategies that hamper clear communications about the size of effects, a lack of use of theory and analysis strategies that facilitate the making of predictions about the future, and generally inadequate estimation and communication of uncertainty. Some of these problems were partially addressed in the empirical component of the thesis, while others became clear only as a result of the empirical engagement itself. In the final chapter of the thesis, I recommended several methodological strategies that psychologists studying the impacts of climate change could use to increase the effectiveness and usefulness of their research. The strategies recommended included using clearly defined categorisations of behaviour rather than psychometric scales, reporting effect sizes in unstandardised form, applying theory that facilitates prediction-making, carefully taking into account the role of time, reporting interval estimates of uncertainty, and addressing as many sources of uncertainty as possible when drawing conclusions.

Performing psychological research concerned with climate change will remain challenging: Predicting how climate change will affect human behaviour in future decades especially so. In this thesis I have identified several important ways in which psychologists' favoured methodological strategies may be problematic when applied to the study of climate change impacts. The ensuing recommendations will hopefully be helpful to psychologists grappling with this important and challenging research area.

8 Consolidated Reference List

Adger, W. N., Pulhin, J. M., Barnett, J., Dabelko, G. D., Hovelsrud, G. K., Levy, M., … Vogel, C. H. (2014). Human security. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, … L. L. White (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 755–791). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Agnew, R. (2011). Dire forecast: A theoretical model of the impact of climate change on crime. *Theoretical Criminology*, *16*(1), 21–42. http://doi.org/10.1177/1362480611416843

Ajdacic-Gross, V., Bopp, M., Ring, M., Gutzwiller, F., & Rossler, W. (2010). Seasonality in suicide – A review and search of new concepts for explaining the heterogeneous phenomena. *Social Science & Medicine*, *71*(4), 657–666. http://doi.org/10.1016/j.socscimed.2010.05.030

Ajdacic-Gross, V., Lauber, C., Sansossio, R., Bopp, M., Eich, D., Gostynski, M., … Rössler, W. (2007). Seasonal associations between weather conditions and suicide—evidence against a classic hypothesis. *American Journal of Epidemiology*, *165*(5), 561–569. http://doi.org/10.1093/aje/kwk034

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *British Medical Journal*, *332*(7549), 1080. http://doi.org/10.1136/bmj.332.7549.1080

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

American Psychological Association. (2012). How does the APA define ‘psychology’? Retrieved 13 February 2012, from http://www.apa.org/support/about/apa/psychology.aspx#answer

Anderegg, W. R. L., Prall, J. W., Harold, J., & Schneider, S. H. (2010). Expert credibility in climate change. *Proceedings of the National Academy of Sciences*, *107*(27), 12107–12109. http://doi.org/10.1073/pnas.1003187107

Anderson, C. A. (1989). Temperature and aggression: Ubiquitous effects of heat on occurrence of human violence. *Psychological Bulletin*, *106*(1), 74–96. http://doi.org/10.1037/0033-2909.106.1.74

Anderson, C. A. (2001). Heat and violence. *Current Directions in Psychological Science*, *10*(1), 33–38. http://doi.org/10.1111/1467-8721.00109

Anderson, C. A., & Anderson, D. C. (1984). Ambient temperature and violent crime: Tests of the linear and curvilinear hypotheses. *Journal of Personality and Social Psychology*, *46*(1), 91–97. http://doi.org/10.1037/0022-3514.46.1.91

Anderson, C. A., & Anderson, K. B. (1996). Violent crime rate studies in philosophical context: A destructive testing approach to heat and southern culture of violence effects. *Journal of Personality and Social Psychology*, *70*(4), 740–756. http://doi.org/10.1037/0022-3514.70.4.740

Anderson, C. A., & Anderson, K. B. (1998). Temperature and aggression: Paradox, controversy, and a (fairly) clear picture. In R. G. Geen & E. D. Donnerstein (Eds.), *Human aggression: Theories, research, and implications for social policy* (pp. 247–298). San Diego, CA: Academic Press.

Anderson, C. A., Anderson, K. B., Dorr, N., DeNeve, K. M., & Flanagan, M. (2000). Temperature and aggression. *Advances in Experimental Social Psychology*, *32*, 63–133. http://doi.org/10.1016/S0065-2601(00)80004-0

Anderson, C. A., Benjamin, A. J., & Bartholow, B. D. (1998). Does the gun pull the trigger? Automatic priming effects of weapon pictures and weapon names. *Psychological Science*, *9*(4), 308–314. http://doi.org/10.1111/1467-9280.00061

Anderson, C. A., Bushman, B. J., & Groom, R. W. (1997). Hot years and serious and deadly assault: Empirical tests of the heat hypothesis. *Journal of Personality and Social Psychology*, *73*(6), 1213–1223. http://doi.org/10.1037/0022-3514.73.6.1213

Anderson, C. A., & DeLisi, M. (2011). Implications of global climate change for violence in developed and developing countries. In J. P. Forgas, A. W. Kruglanski, & K. D. Williams (Eds.), *The psychology of social conflict and aggression.* (pp. 249–265). New York, NY: Psychology Press.

Anderson, C. A., Deuser, W. E., & DeNeve, K. M. (1995). Hot temperatures, hostile affect, hostile cognition, and arousal: Tests of a general model of affective aggression. *Personality and Social Psychology Bulletin*, *21*(5), 434–448. http://doi.org/10.1177/0146167295215002

Archer, D. (2009). *The long thaw: How humans are changing the next 100,000 years of Earth's climate*. Princeton, NJ: Princeton University Press.

Aronow, P. M., & Samii, C. (in press). Does regression produce representative estimates of causal effects? *American Journal of Political Science*. Retrieved from http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2224964

Arrhenius, S. (1896). On the influence of carbonic acid in the air upon the temperature of the ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *41*(251), 237–276. http://doi.org/10.1080/14786449608620846

Arrhenius, S. (1908). *Worlds in the making: The evolution of the universe*. New York, NY: Harper.

Australian Bureau of Meteorology. (n.d.). Southern Oscillation Index (SOI) Archives. Retrieved 15 January 2015, from http://www.bom.gov.au/climate/current/soihtm1.shtml

Bååth, R. (2014). Probable points and credible intervals, part 1: Graphical summaries. Retrieved 19 February 2015, from http://www.sumsar.net/blog/2014/10/probable-points-and-credible-intervals-part-one/

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, *100*(3), 603–617. http://doi.org/10.1348/000712608X377117

Baguley, T. (2010). When correlations go bad. *The Psychologist*, *23*(2), 122–123.

Bammer, G., & Smithson, M. (2008). *Uncertainty and risk: Multidisciplinary perspectives*. London, United Kingdom: Earthscan.

Barker, A., Hawton, K., Fagg, J., & Jennison, C. (1994). Seasonal and weather factors in parasuicide. *The British Journal of Psychiatry*, *165*(3), 375–380. http://doi.org/10.1192/bjp.165.3.375

Baron, R. A., & Bell, P. A. (1975). Aggression and heat: Mediating effects of prior provocation and exposure to an aggressive model. *Journal of Personality and Social Psychology*, *31*(5), 825–832. http://doi.org/10.1037/h0076647

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version 1.0–5). Retrieved from http://cran.r-project.org/web/packages/lme4/index.html

Baum, A., Fleming, I., Israel, A., & O'Keeffe, M. K. (1992). Symptoms of chronic stress following a natural disaster and discovery of a human-made hazard. *Environment and Behavior*, *24*(3), 347–365. http://doi.org/10.1177/0013916592243004

Bayes, T., & Price, R. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418. http://doi.org/10.1098/rstl.1763.0053

Bell, P. A., & Baron, R. A. (1976). Aggression and heat: The mediating role of negative affect. *Journal of Applied Social Psychology*, *6*(1), 18–30. http://doi.org/10.1111/j.1559-1816.1976.tb01308.x

Bell, P. A., & Fusco, M. E. (1989). Heat and violence in the Dallas field data: Linearity, curvilinearity, and heteroscedasticity. *Journal of Applied Social Psychology*, *19*(17), 1479–1482. http://doi.org/10.1111/j.1559-1816.1989.tb01459.x

Bennetts, R., & Seabrook, R. (2008). Retail sales of alcohol and the risk of being a victim of assault. *PLoS Medicine*, *5*(5). http://doi.org/10.1371/journal.pmed.0050108

Berk, M., Dodd, S., & Henry, M. (2006). The effect of macroeconomic variables on suicide. *Psychological Medicine*, *36*(2), 181–189. http://doi.org/10.1017/S0033291705006665

Berk, R. A. (2004). *Regression analysis: A constructive critique*. Thousand Oaks, CA: Sage.

Berry, H. L., Bowen, K., & Kjellstrom, T. (2010). Climate change and mental health: A causal pathways framework. *International Journal of Public Health*, *55*(2), 123–132. http://doi.org/10.1007/s00038-009-0112-0

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135. http://doi.org/10.1016/j.tree.2008.10.008

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge, UK: Cambridge University Press.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. http://doi.org/10.1037/0033-295X.110.2.203

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. http://doi.org/10.1037/0033-295X.111.4.1061

Budescu, D. V., Broomell, S., & Por, H.-H. (2009). Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. *Psychological Science*, *20*(3), 299–308. http://doi.org/10.1111/j.1467-9280.2009.02284.x

Bushman, B. J., & Cooper, H. M. (1990). Effects of alcohol on human aggression: An integrative research review. *Psychological Bulletin*, *107*(3), 341–354. http://doi.org/10.1037/0033-2909.107.3.341

Bushman, B. J., Wang, M. C., & Anderson, C. A. (2005). Is the curve relating temperature to aggression linear or curvilinear? Assaults and temperature in Minneapolis reexamined. *Journal of Personality and Social Psychology*, *89*(1), 62–66. http://doi.org/10.1037/0022-3514.89.1.62

Buttice, M. K., & Highton, B. (2013). How does multilevel regression and poststratification perform with conventional national surveys? *Political Analysis*, *21*(4), 449–467. http://doi.org/10.1093/pan/mpt017

Cacioppo, J. T., Semin, G. R., & Berntson, G. G. (2004). Realism, instrumentalism, and scientific symbiosis: Psychological theory as a search for truth and the discovery of solutions. *American Psychologist*, *59*(4), 214–223. http://doi.org/10.1037/0003-066X.59.4.214

Ceccato, V. (2005). Homicide in São Paulo, Brazil: Assessing spatial-temporal and weather variations. *Journal of Environmental Psychology*, *25*(3), 307–321. http://doi.org/10.1016/j.jenvp.2005.07.002

Centre for Disease Control and Prevention. (2013). International classification of diseases, ninth revision, clinical modification (ICD-9-CM). Retrieved from http://www.cdc.gov/nchs/icd/icd9cm.htm

Chang, S.-S., Gunnell, D., Sterne, J. A. C., Lu, T.-H., & Cheng, A. T. A. (2009). Was the economic crisis 1997–1998 responsible for rising suicide rates in East/Southeast Asia? A time–trend analysis for Japan, Hong Kong, South Korea, Taiwan, Singapore and Thailand. *Social Science & Medicine*, *68*(7), 1322–1331. http://doi.org/10.1016/j.socscimed.2009.01.010

Chatfield, C. (1993). Calculating interval forecasts. *Journal of Business & Economic Statistics*, *11*(2), 121–135. http://doi.org/10.2307/1391361

Chew, K. S. Y., & McCleary, R. (1995). The spring peak in suicides: A cross-national analysis. *Social Science & Medicine*, *40*(2), 223–230. http://doi.org/10.1016/0277-9536(94)E0070-9

Chow, S. L. (1988). Significance test or effect size? *Psychological Bulletin*, *103*(1), 105–110. http://doi.org/10.1037/0033-2909.103.1.105

Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., … Whetton, P. (2007). Regional climate projections. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, … H. L. Miller (Eds.), *Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, *7*(3), 249–253. http://doi.org/10.1177/014662168300700301

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist*, *49*(12), 997–1003. http://doi.org/10.1037/0003-066X.49.12.997

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American Sociological Review*, *44*(4), 588–608. http://doi.org/10.2307/2094589

Cohn, E. G., & Rotton, J. (1997). Assault as a function of time and temperature: A moderator-variable time-series analysis. *Journal of Personality and Social Psychology*, *72*(6), 1322–1334. http://doi.org/10.1037/0022-3514.72.6.1322

Cohn, E. G., & Rotton, J. (2000). Weather, seasonal trends and property crimes in Minneapolis, 1987–1988. A moderator-variable time-series analysis of routine activities. *Journal of Environmental Psychology*, *20*(3), 257–272. http://doi.org/10.1006/jevp.1999.0157

Cohn, E. G., & Rotton, J. (2005). The curve is still out there: A reply to Bushman, Wang, and Anderson's (2005) 'Is the curve relating temperature to aggression linear or curvilinear?' *Journal of Personality and Social Psychology*, *89*(1), 67–70. http://doi.org/10.1037/0022-3514.89.1.67

Cohn, E. G., Rotton, J., Peterson, A. G., & Tarr, D. B. (2004). Temperature, city size, and the Southern subculture of violence: Support for social escape/avoidance (SEA) theory. *Journal of Applied Social Psychology*, *34*(8), 1652–1674. http://doi.org/10.1111/j.1559-1816.2004.tb02792.x

Collins, M., Knutti, R., Arblaster, J. M., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., … Wehner, M. (2013). Long-term climate change: Projections, commitments and irreversibility. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, … P. M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., … Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, *8*(2). http://doi.org/10.1088/1748-9326/8/2/024024

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121–136. http://doi.org/10.1080/00223890802634175

Cozby, P. C., & Bates, S. C. (2015). *Methods in behavioral research* (12th ed.). New York, NY: McGraw-Hill.

Crimes Act. (1961). Retrieved 25 September 2012, from http://www.legislation.govt.nz/act/public/1961/0043/latest/DLM327382.html

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. http://doi.org/10.1037/h0040957

Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.

Cumming, G., Fidler, F., Leonard, M., Kalinowski, P., Christiansen, A., Kleinig, A., … Wilson, S. (2007). Statistical reform in psychology: Is anything changing? *Psychological Science*, *18*(3), 230–232. http://doi.org/10.1111/j.1467-9280.2007.01881.x

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., … others. (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, *364*(6434), 218–220. http://doi.org/10.1038/364218a0

Dattilo, A. M., & Kris-Etherton, P. M. (1992). Effects of weight reduction on blood lipids and lipoproteins: A meta-analysis. *The American Journal of Clinical Nutrition*, *56*(2), 320–328.

Deisenhammer, E. A. (2003). Weather and suicide: The present state of knowledge on the association of meteorological factors with suicidal behaviour. *Acta Psychiatrica Scandinavica*, *108*(6), 402–409. http://doi.org/10.1046/j.0001-690X.2003.00209.x

Deisenhammer, E. A., Kemmler, G., & Parson, P. (2003). Association of meteorological factors with suicide. *Acta Psychiatrica Scandinavica*, *108*(6), 455–459. http://doi.org/10.1046/j.0001-690X.2003.00219.x

Dell, M., Jones, B. F., & Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, *4*(3), 66–95. http://doi.org/10.1257/mac.4.3.66

Dixon, P. G., & Kalkstein, A. J. (2009). Climate suicide relationships: A research problem in need of geographic methods and cross-disciplinary perspectives. *Geography Compass*, *3*(6), 1961–1974. http://doi.org/10.1111/j.1749-8198.2009.00286.x

Dixon, P. G., Sinyor, M., Schaffer, A., Levitt, A., Haney, C. R., Ellis, K. N., & Sheridan, S. C. (2014). Association of weekly suicide rates with temperature anomalies in two different climate types. *International Journal of Environmental Research and Public Health*, *11*(11), 11627–11644. http://doi.org/10.3390/ijerph111111627

Doganay, Z., Sunter, A. T., Guz, H., Ozkan, A., Altintop, L., Kati, C., … Guven, H. (2003). Climatic and diurnal variation in suicide attempts in the ED. *The American Journal of Emergency Medicine*, *21*(4), 271–275. http://doi.org/10.1016/S0735-6757(03)00039-1

Doherty, T. J., & Clayton, S. (2011). The psychological impacts of global climate change. *American Psychologist*, *66*(4), 265–276. http://doi.org/10.1037/a0023141

Durkheim, É. (1897). *Le Suicide*. Paris, France: Léo.

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*(8746), 867–872. http://doi.org/10.1016/0140-6736(91)90201-Y

Efron, B. (1986). Why isn't everyone a Bayesian? *The American Statistician*, *40*(1), 1–5. http://doi.org/10.1080/00031305.1986.10475342

Federal Bureau of Investigation. (2010). Crime reported by Dallas Police Dept, Texas: Crime rate per 100,000 population, violent crime 1990–2010. Retrieved 7 February 2013, from http://www.ucrdatatool.gov/Search/Crime/Local/RunCrimeJurisbyJurisLarge.cfm

Fisher, R. A. (1921). On the ‘probable error’ of a coefficient of correlation deduced from a small sample. *Metron*, *1*, 3–32.

Fox, J. (1997). *Applied regression analysis, linear models, and related methods*. Thousand Oaks, CA: Sage.

Fraedrich, K., Aigner, A. A., Kirk, E., & Lunkeit, F. (2006). General circulation models of the atmosphere. In A. Scott (Ed.), *Encyclopedia of nonlinear science* (pp. 359–361). London, United Kingdom: Routledge.

Frich, P., Alexander, L. V., Della-Marta, P., Gleason, B., Haylock, M., Klein Tank, A. M., & Peterson, T. (2002). Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, *19*(3), 193–212. http://doi.org/10.3354/cr019193

Friedman, S. M., Dunwoody, S., & Rogers, C. L. (1999). *Communicating uncertainty: Media coverage of new and controversial science*. New York, NY: Routledge.

Gamble, J. L., & Hess, J. J. (2012). Temperature and violent crime in Dallas, Texas: Relationships and implications of climate change. *Western Journal of Emergency Medicine*, *13*(3), 239–246. http://doi.org/10.5811/westjem.2012.3.11746

Gardner, M. J., & Altman, D. G. (1986). Confidence intervals rather than P values: Estimation rather than hypothesis testing. *British Medical Journal*, *292*(6522), 746–750. http://doi.org/10.1136/bmj.292.6522.746

Gelman, A. (2004). Treatment effects in before-after data. In A. Gelman & X.-L. Meng (Eds.), *Applied Bayesian modeling and causal inference from an incomplete data perspective* (pp. 195–202). London, United Kingdom: Wiley.

Gelman, A. (2008). Objections to Bayesian statistics. *Bayesian Analysis*, *3*(3), 445–449. http://doi.org/10.1214/08-BA318

Gelman, A. (2015). The connection between varying treatment effects and the crisis of unreplicable research: A Bayesian perspective. *Journal of Management*, *41*(2), 632–643. http://doi.org/10.1177/0149206314525208

Gelman, A., & Carlin, B. P. (2001). Poststratification and weighting adjustments. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 288–302). New York, NY: Wiley.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, United Kingdom: Cambridge University Press.

Gelman, A., & Huang, Z. (2008). Estimating incumbency advantage and its variation, as an example of a before–after study. *Journal of the American Statistical Association*, *103*(482), 437–446. http://doi.org/10.1198/016214507000000626

Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no 'fishing expedition' or 'p-hacking' and the research hypothesis was posited ahead of time*. Retrieved from http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

Gelman, A., Su, Y.-S., Yajima, M., Hill, J., Pittau, M. G., Kerman, J., … Dorie, V. (2014). arm: Data analysis using regression and multilevel/hierarchical models (Version 1.7–03). Retrieved from http://cran.r-project.org/web/packages/arm/

Gifford, R. (2008). Psychology's essential role in alleviating the impacts of climate change. *Canadian Psychology*, *49*(4), 273–280. http://doi.org/10.1037/a0013234

Gifford, R. (2011). The dragons of inaction: Psychological barriers that limit climate change mitigation and adaptation. *American Psychologist*, *66*(4), 290–302. http://doi.org/10.1037/a0023566

Gigerenzer, G., Krauss, S., & Vitouch, O. (2004). The null ritual: What you always wanted to know about significance testing but were afraid to ask. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 391–408). Thousand Oaks, CA: Sage.

Gillis, J. (2013, May 10). Carbon dioxide level passes long-feared milestone. *The New York Times*. Retrieved from http://www.nytimes.com/2013/05/11/science/earth/carbon-dioxide-level-passes-long-feared-milestone.html

Gill, J. (1999). The insignificance of null hypothesis significance testing. *Political Research Quarterly*, *52*(3), 647–674. http://doi.org/10.1177/106591299905200309

Gill, J. (2015). *Bayesian methods: A social and behavioral sciences approach* (3rd ed.). Boca Raton, FL: CRC Press.

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, *9*(3), 330–338. http://doi.org/10.1016/j.sapharm.2012.04.004

Goodman, S. (2008). A dirty dozen: Twelve *P*-value misconceptions. *Seminars in Hematology*, *45*(3), 135–140. http://doi.org/10.1053/j.seminhematol.2008.04.003

Greenland, S., & Poole, C. (2013). Living with *P* values: Resurrecting a Bayesian perspective on frequentist statistics. *Epidemiology*, *24*(1), 62–68. http://doi.org/10.1097/EDE.0b013e3182785741

Grove, O., & Lynge, J. (1979). Suicide and attempted suicide in Greenland. *Acta Psychiatrica Scandinavica*, *60*(4), 375–391. http://doi.org/10.1111/j.1600-0447.1979.tb00286.x

Grove, W. M., & Lloyd, M. (2006). Meehl's contribution to clinical versus statistical prediction. *Journal of Abnormal Psychology*, *115*(2), 192–194. http://doi.org/10.1037/0021-843X.115.2.192

Gupta, A. K. (2004). Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Current Science*, *87*(1), 54–59.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*(2). Retrieved from http://www.jstatsoft.org/v33/i02

Haig, B. D. (2005). An abductive theory of scientific method. *Psychological Methods*, *10*(4), 371–388. http://doi.org/10.1037/1082-989X.10.4.371

Hansen, J., & Hansen, A. (2008). Dear Michelle and Barack Obama. Retrieved from http://www.columbia.edu/~jeh1/mailings/2008/20081229_DearMichelleAndBarack.pdf

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, *48*(4). http://doi.org/10.1029/2010RG000345

Harries, K. D., & Stadler, S. J. (1983). Determinism revisited: Assault and heat stress in Dallas, 1980. *Environment and Behavior*, *15*(2), 235–256. http://doi.org/10.1177/0013916583152006

Harries, K. D., & Stadler, S. J. (1988). Heat and violence: New findings from Dallas field data, 1980–1981. *Journal of Applied Social Psychology*, *18*(2), 129–138. http://doi.org/10.1111/j.1559-1816.1988.tb00010.x

Hays, R. D., Marshall, G. N., Wang, E. Y. ., & Sherbourne, C. D. (1994). Four-year cross-lagged associations between physical and mental health in the Medical Outcomes Study. *Journal of Consulting and Clinical Psychology*, *62*(3), 441–449. http://doi.org/10.1037/0022-006X.62.3.441

Helama, S., Holopainen, J., & Partonen, T. (2013). Temperature-associated suicide mortality: Contrasting roles of climatic warming and the suicide prevention program in Finland. *Environmental Health and Preventive Medicine*, *18*(5), 349–355. http://doi.org/10.1007/s12199-013-0329-7

Held, I. M., & Soden, B. J. (2000). Water vapor feedback and global warming. *Annual Review of Energy and the Environment*, *25*(1), 441–475. http://doi.org/10.1146/annurev.energy.25.1.441

Hesterberg, T., Moore, D. S., Monaghan, S., Clipson, A., & Epstein, R. (2002). Bootstrap methods and permutation tests. In D. S. Moore & G. P. McCabe (Eds.), *Introduction to the practice of statistics* (4th ed.). New York, NY: Freeman.

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge, United Kingdom: Cambridge University Press.

Hiltunen, L., Ruuhela, R., Ostamo, A., Lönnqvist, J., Suominen, K., & Partonen, T. (2012). Atmospheric pressure and suicide attempts in Helsinki, Finland. *International Journal of Biometeorology*, *56*(6), 1045–1053. http://doi.org/10.1007/s00484-011-0518-2

Hipp, J. R., Bauer, D. J., Curran, P. J., & Bollen, K. A. (2003). Crimes of opportunity or crimes of emotion? Testing two explanations of seasonal change in crime. *Social Forces*, *82*(4), 1333–1372. http://doi.org/10.1353/sof.2004.0074

Hoekstra, R., Kiers, H. A. L., & Johnson, A. (2012). Are assumptions of well-known statistical techniques checked, and why (not)? *Frontiers in Psychology*, *3*. http://doi.org/10.3389/fpsyg.2012.00137

Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E.-J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, *21*(5), 1157–1164. http://doi.org/10.3758/s13423-013-0572-3

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*(396), 945–960. http://doi.org/10.1080/01621459.1986.10478354

Holland, P. W., & Rubin, D. B. (1988). Causal inference in retrospective studies. *Evaluation Review*, *12*(3), 203–231. http://doi.org/10.1177/0193841X8801200301

Holopainen, J., Helama, S., Björkenstam, C., & Partonen, T. (2013). Variation and seasonal patterns of suicide mortality in Finland and Sweden since the 1750s. *Environmental Health and Preventive Medicine*, *18*(6), 494–501. http://doi.org/10.1007/s12199-013-0348-4

Holopainen, J., Helama, S., & Partonen, T. (2014). Does diurnal temperature range influence seasonal suicide mortality? Assessment of daily data of the Helsinki metropolitan area from 1973 to 2010. *International Journal of Biometeorology*, *58*(6), 1039–1045. http://doi.org/10.1007/s00484-013-0689-0

Horrocks, J., & Menclova, A. K. (2011). The effects of weather on crime. *New Zealand Economic Papers*, *45*(3), 231–254. http://doi.org/10.1080/00779954.2011.572544

Hsiang, S. M. (2013). Visually-weighted regression. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2265501

Hsiang, S. M., Burke, M., & Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, *341*(6151), 1212–1226. http://doi.org/10.1126/science.1235367

Hurvich, C. M., Simonoff, J. S., & Tsai, C.-L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(2), 271–293. http://doi.org/10.1111/1467-9868.00125

International Bureau of Weights and Measures. (2006). *The international system of units (SI)* (8th ed.). Retrieved from http://www.bipm.org/utils/common/pdf/si_brochure_8_en.pdf

IPCC. (2013a). *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY: Cambridge University Press.

IPCC. (2013b). Summary for policymakers. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, … P. M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 1–29). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

IPCC. (2014a). *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: Author.

IPCC. (2014b). Summary for policymakers. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, … L. L. White (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Contribution of working group II to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 1–32). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

iTouchMap.com. (2013). Latitude and longitude of a point. Retrieved 1 October 2013, from http://itouchmap.com/latlong.html

Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills: A practical guide for social scientists*. New York, NY: Guilford Press.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. http://doi.org/10.1037/0022-006X.59.1.12

Jessen, G., Andersen, K., Arensman, E., Bille-Brahe, U., Crepet, P., De Leo, D., … Wasserman, D. (1999). Temporal fluctuations and seasonality in attempted suicide in Europe: Findings from the WHO/EURO multicentre study on parasuicide. *Archives of Suicide Research*, *5*(1), 57–69. http://doi.org/10.1080/13811119908258315

Jessen, G., Steffensen, P., & Jensen, B. F. (1998). Seasons and meteorological factors in suicidal behaviour: Findings and methodological considerations from a Danish study. *Archives of Suicide Research*, *4*(3), 263–280. http://doi.org/10.1080/13811119808258301

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). New York, NY: Guilford Press.

Keeling, C. D. (1960). The concentration and isotopic abundances of carbon dioxide in the atmosphere. *Tellus*, *12*(2), 200–203. http://doi.org/10.1111/j.2153-3490.1960.tb01300.x

Kim, J., & Seltzer, M. (2011). Examining heterogeneity in residual variance to detect differential response to treatments. *Psychological Methods*, *16*(2), 192–208. http://doi.org/10.1037/a0022656

Kim, Y., Kim, H., & Kim, D.-S. (2011). Association between daily environmental temperature and suicide mortality in Korea (2001–2005). *Psychiatry Research*, *186*(2–3), 390–396. http://doi.org/10.1016/j.psychres.2010.08.006

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*(5), 746–759. http://doi.org/10.1177/0013164496056005002

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Kruschke, J. K. (2010a). Bayesian data analysis. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(5), 658–676. http://doi.org/10.1002/wcs.72

Kruschke, J. K. (2010b). *Doing Bayesian data analysis: A tutorial with R and BUGS* (1st ed.). Burlington, MA: Academic Press.

Kyngdon, A. (2011). Plausible measurement analogies to some psychometric models of test performance. *British Journal of Mathematical and Statistical Psychology*, *64*(3), 478–497. http://doi.org/10.1348/2044-8317.002004

Leffingwell, A. (1892). *Illegitimacy, and the influence of seasons upon conduct: Two studies in demography*. London, United Kingdom: Sonnenschein.

Leick, G. (2002). *Mesopotamia: The invention of the city*. London, United Kingdom: Penguin.

Lester, D. (1999). Climatic data and national suicide and homicide rates. *Perceptual and Motor Skills*, *89*(3), 1036. http://doi.org/10.2466/pms.1999.89.3.1036

Lewin, R. (1987). Africa: Cradle of modern humans. *Science*, *237*(4820), 1292–1295. http://doi.org/10.1126/science.3114877

Lewis, J. (1990). The vulnerability of small island states to sea level rise: The need for holistic strategies. *Disasters*, *14*(3), 241–249. http://doi.org/10.1111/j.1467-7717.1990.tb01066.x

LGNZ. (2015). Council maps and websites. Retrieved 5 February 2015, from http://www.lgnz.co.nz/home/nzs-local-government/new-zealands-councils/

Likhvar, V., Honda, Y., & Ono, M. (2011). Relation between temperature and suicide mortality in Japan in the presence of other confounding factors using time-series analysis with a semiparametric approach. *Environmental Health and Preventive Medicine*, *16*(1), 36–43. http://doi.org/10.1007/s12199-010-0163-0

Lind, J. T., & Mehlum, H. (2010). With or without U? The appropriate test for a U-shaped relationship. *Oxford Bulletin of Economics and Statistics*, *72*(1), 109–118. http://doi.org/10.1111/j.1468-0084.2009.00569.x

Lindsay, J. J., & Anderson, C. A. (2000). From antecedent conditions to violent actions: A general affective aggression model. *Personality and Social Psychology Bulletin*, *26*(5), 533–547. http://doi.org/10.1177/0146167200267002

Li, X., Harbottle, G., Zhang, J., & Wang, C. (2003). The earliest writing? Sign use in the seventh millennium BC at Jiahu, Henan Province, China. *Antiquity*, *77*(295), 31–44. http://doi.org/10.1017/S0003598X00061329

Luce, R. D. (1999). Where is mathematical modeling in psychology headed? *Theory & Psychology*, *9*(6), 723–737. http://doi.org/10.1177/0959354399096001

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, *1*(1), 1–27. http://doi.org/10.1016/0022-2496(64)90015-X

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*, 151–169. http://doi.org/annurev.publhealth.23.100901.140546

Lyons, L. (2013). Bayes and frequentism: A particle physicist's perspective. *Contemporary Physics*, *54*(1), 1–16. http://doi.org/10.1080/00107514.2012.756312

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. http://doi.org/10.1037/1082-989X.7.1.19

Mackintosh, L. (2001). Overview of New Zealand climate. Retrieved 10 January 2012, from http://www.niwa.co.nz/education-and-training/schools/resources/climate/overview

MacPherson, L. (2014). National population estimates: At 30 June 2014. Retrieved 29 January 2015, from http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationEstimates_HOTPAt30Jun14.aspx

Madge, N., Hewitt, A., Hawton, K., Wilde, E. J. de, Corcoran, P., Fekete, S., … Ystgaard, M. (2008). Deliberate self-harm within an international community sample of young people: Comparative findings from the Child & Adolescent Self-harm in Europe (CASE) Study. *Journal of Child Psychology and Psychiatry*, *49*(6), 667–677. http://doi.org/10.1111/j.1469-7610.2008.01879.x

Maes, M., Meyer, F., Thompson, P., Peeters, D., & Cosyns, P. (1994). Synchronized annual rhythms in violent suicide rate, ambient temperature and the light-dark span. *Acta Psychiatrica Scandinavica*, *90*(5), 391–396. http://doi.org/10.1111/j.1600-0447.1994.tb01612.x

Maes, M., Scharpe, S., Verkerk, R., D'Hondt, P., Peeters, D., Cosyns, P., … Neels, H. (1995). Seasonal variation in plasma L-tryptophan availability in healthy volunteers: Relationships to violent suicide occurrence. *Archives of General Psychiatry*, *52*(11), 937–946. http://doi.org/10.1001/archpsyc.1995.03950230051008

Mann, M. E. (2012). *The hockey stick and the climate wars: Dispatches from the front lines*. New York, NY: Columbia University Press.

Mann, P. S. (1994). *Statistics for business and economics*. Toronto, Canada: Wiley.

Mares, D. (2013). Climate change and crime: Monthly temperature and precipitation anomalies and crime rates in St. Louis, MO 1990–2009. *Crime, Law and Social Change*, *59*(2), 185–208. http://doi.org/10.1007/s10611-013-9411-8

Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*, *137*(5), 856–879. http://doi.org/10.1037/a0023678

Masterton, G. (1991). Monthly and seasonal variation in parasuicide. A sex difference. *The British Journal of Psychiatry*, *158*(2), 155–157. http://doi.org/10.1192/bjp.158.2.155

Mastrandrea, M., Mach, K., Plattner, G.-K., Edenhofer, O., Stocker, T., Field, C., … Matschoss, P. (2011). The IPCC AR5 guidance note on consistent treatment of uncertainties: A common approach across the working groups. *Climatic Change*, *108*(4), 675–691. http://doi.org/10.1007/s10584-011-0178-6

Mayr, E., & Bock, W. J. (2002). Classifications and other ordering systems. *Journal of Zoological Systematics and Evolutionary Research*, *40*(4), 169–194. http://doi.org/10.1046/j.1439-0469.2002.00211.x

McCandless, L. C., Gustafson, P., & Levy, A. (2007). Bayesian sensitivity analysis for unmeasured confounding in observational studies. *Statistics in Medicine*, *26*(11), 2331–2347. http://doi.org/10.1002/sim.2711

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, *107*(2), 247–255. http://doi.org/10.1037/0033-2909.107.2.247

McFarlane, A. C., & Van Hooff, M. (2009). Impact of childhood exposure to a natural disaster on adult mental health: 20-year longitudinal follow-up study. *The British Journal of Psychiatry*, *195*(2), 142–148. http://doi.org/10.1192/bjp.bp.108.054270

McMichael, A. J., Woodruff, R. E., & Hales, S. (2006). Climate change and human health: Present and future risks. *The Lancet*, *367*(9513), 859–869. http://doi.org/16/S0140-6736(06)68079-3

McNamee, R. (2003). Confounding and confounders. *Occupational and Environmental Medicine*, *60*(3), 227–234. http://doi.org/10.1136/oem.60.3.227

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749. http://doi.org/10.1037/0003-066X.50.9.741

Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin*, *100*(3), 398–407. http://doi.org/10.1037/0033-2909.100.3.398

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, *88*(3), 355–383. http://doi.org/10.1111/j.2044-8295.1997.tb02641.x

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept*. Cambridge, United Kingdom: Cambridge University Press.

Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory & Psychology*, *10*(5), 639–667. http://doi.org/10.1177/0959354300105004

Michell, J. (2008). Is psychometrics pathological science? *Measurement*, *6*(1–2), 7–24. http://doi.org/10.1080/15366360802035489

Michell, J. (2009). Invalidity in validity. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 111–133). Charlotte, NC: Information Age.

Michell, J. (2011). Qualitative research meets the ghost of Pythagoras. *Theory & Psychology*, *21*(2), 241–259. http://doi.org/10.1177/0959354310391351

Michell, J. (2012). 'The constantly recurring argument': Inferring quantity from order. *Theory & Psychology*, *22*(3), 255–271. http://doi.org/10.1177/0959354311434656

Michell, J., & Ernst, C. (1996). The axioms of quantity and the theory of measurement: Translated from part I of Otto Hölder's German text 'Die axiome der quantität und die lehre vom mass'. *Journal of Mathematical Psychology*, *40*(3), 235–252. http://doi.org/10.1006/jmps.1996.0023

Ministry for the Environment. (2008). *Climate change effects and impacts assessment: A guidance manual for local government in New Zealand* (2nd ed.). Wellington, New Zealand: Author.

Ministry of Health. (2012). *Suicide facts: Deaths and intentional self-harm hospitalisations 2010*. Retrieved from http://www.health.govt.nz/publication/suicide-facts-deaths-and-intentional-self-harm-hospitalisations-2010

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2001). *Introduction to linear regression analysis* (3rd ed.). New York, NY: Wiley.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2015). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*. Advance online publication. http://doi.org/10.3758/s13423-015-0947-8

Morselli, E. A. (1882). *Suicide: An essay on comparative moral statistics*. New York, NY: Appleton.

Moser, S. C. (2010). Communicating climate change: History, challenges, process and future directions. *Wiley Interdisciplinary Reviews: Climate Change*, *1*(1), 31–53. http://doi.org/10.1002/wcc.11

Moss, R. H., Babiker, M., Brinkman, S., Calvo, E., Carter, T., Edmonds, J. A., … Zurek, M. (2008). *Towards new scenarios for analysis of emissions, climate* *change, impacts, and response strategies: Technical summary*. Geneva, Switzerland: Intergovernmental Panel on Climate Change.

Müller, H., Biermann, T., Renk, S., Reulbach, U., Ströbel, A., Kornhuber, J., & Sperling, W. (2011). Higher environmental temperature and global radiation are correlated with increasing suicidality—a localized data analysis. *Chronobiology International*, *28*(10), 949–957. http://doi.org/10.3109/07420528.2011.618418

Myung, I. J., & Pitt, M. A. (2002). Mathematical modeling. In H. Pashler & J. Wixted (Eds.), *Stevens' Handbook of Experimental Psychology: Methodology in Experimental Psychology* (3rd ed., Vol. 4, pp. 429–460). Hoboken, NJ: Wiley.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biological Reviews*, *82*(4), 591–605. http://doi.org/10.1111/j.1469-185X.2007.00027.x

NASA. (n.d.). Navier-Stokes equations. Retrieved from http://www.grc.nasa.gov/WWW/k-12/airplane/nseqs.html

National Aeronautics and Space Administration. (n.d.). GISS surface temperature analysis: Analysis graphs and plots. Retrieved 3 March 2015, from http://data.giss.nasa.gov/gistemp/graphs_v3/

National Climatic Data Center. (n.d.). Paleoclimatology data. Retrieved 4 March 2015, from http://www.ncdc.noaa.gov/data-access/paleoclimatology-data

National Oceanic and Atmospheric Administration. (n.d.). National Climatic Data Center. Retrieved 4 February 2013, from http://www.ncdc.noaa.gov/

Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical and Physical Sciences*, *236*(767), 333–380. http://doi.org/10.1098/rsta.1937.0005

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*(2), 241–301. http://doi.org/10.1037/1082-989X.5.2.241

Nie, Z., & Racine, J. S. (2012). The crs package: Nonparametric regression splines for continuous and categorical predictors. *The R Journal*, *4*(2), 48–56.

NIWA. (n.d.-a). CliFlo national climate database. Retrieved 3 October 2013, from http://cliflo.niwa.co.nz/

NIWA. (n.d.-b). Virtual climate station data and products. Retrieved 10 October 2012, from http://www.niwa.co.nz/our-science/climate/our-services/virtual-climate-stations

Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981–2001. *Psychiatry: Interpersonal and Biological Processes*, *65*(3), 207–239. http://doi.org/10.1521/psyc.65.3.207.20173

Oravecz, R., Czigler, B., & Moore, M. (2006). The transformation of suicide fluctuation in Slovenia. *Archives of Suicide Research*, *10*(1), 69–76. http://doi.org/10.1080/13811110500318471

Page, L. A., Hajat, S., & Kovats, R. (2007). Relationship between daily suicide counts and temperature in England and Wales. *The British Journal of Psychiatry*, *191*(2), 106–112. http://doi.org/10.1192/bjp.bp.106.031948

Page, L. A., & Howard, L. M. (2010). The impact of climate change on mental health (but will mental health be discussed at Copenhagen?). *Psychological Medicine*, *12*(2), 177–180. http://doi.org/10.1017/S0033291709992169

Pavot, W. (2008). The assessment of subjective well-being: Successes and shortfalls. In M. Eid & R. J. Larsen (Eds.), *The science of subjective well-being* (pp. 124–140). New York, NY: Guilford Press.

Pawlik, K. (1992). The psychology of global environmental change: Some basic data and an agenda for cooperative international research. *International Journal of Psychology*, *26*(5), 547–563. http://doi.org/10.1080/00207599108247143

Perez, B. M. (2011). The aura: A brief review. *Skeptical Inquirer*, *35*(1). Retrieved from http://www.csicop.org/si/show/the_aura_a_brief_review/

Pidgeon, N., & Fischhoff, B. (2011). The role of social and decision sciences in communicating uncertain climate risks. *Nature Climate Change*, *1*(1), 35–41. http://doi.org/10.1038/nclimate1080

Pinker, S. (2011). *The better angels of our nature: Why violence has declined*. London, United Kingdom: Penguin.

Planton, S. (2013). Annex III: Glossary. In T. F. Stocker, D. Qin, G.-. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, … P. M. Midgley (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 1447–1465). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Preti, A. (1997). The influence of seasonal change on suicidal behaviour in Italy. *Journal of Affective Disorders*, *44*(2–3), 123–130. http://doi.org/10.1016/S0165-0327(97)00035-9

Preti, A. (1998). The influence of climate on suicidal behaviour in Italy. *Psychiatry Research*, *78*(1–2), 9–19. http://doi.org/10.1016/S0165-1781(97)00154-6

Preti, A., Lentini, G., & Maugeri, M. (2007). Global warming possibly linked to an enhanced risk of suicide: Data from Italy, 1974–2003. *Journal of Affective Disorders*, *102*(1–3), 19–25. http://doi.org/10.1016/j.jad.2006.12.003

Price, P. (2012). *Research methods in psychology*. Washington, DC: Saylor Foundation.

Qi, X., Hu, W., Mengersen, K., & Tong, S. (2014). Socio-environmental drivers and suicide in Australia: Bayesian spatial analysis. *BMC Public Health*, *14*(681). http://doi.org/10.1186/1471-2458-14-681

Qi, X., Tong, S., & Hu, W. (2009). Preliminary spatiotemporal analysis of the association between socio-environmental factors and suicide. *Environmental Health*, *8*(46). http://doi.org/10.1186/1476-069X-8-46

Rachlinski, J. J. (2000). The psychology of global climate change. *University of Illinois Law Review*, *2000*(1), 299–319.

Ramachandran, K. M., & Tsokos, C. P. (2009). *Mathematical statistics with applications*. Burlington, MA: Elsevier.

Rameesh, R. (2008, November 10). Paradise almost lost: Maldives seek to buy a new homeland. *The Guardian*. Retrieved from http://www.theguardian.com/environment/2008/nov/10/maldives-climate-change

Ray, J. G., Moineddin, R., Bell, C. M., Thiruchelvam, D., Creatore, M. I., Gozdyra, P., … Redelmeier, D. A. (2008). Alcohol sales and risk of serious assault. *PLoS Medicine*, *5*(5). http://doi.org/10.1371/journal.pmed.0050104

R Core Team. (2013). R: A language and environment for statistical computing (Version 3.0.2). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reisinger, A., Kitching, R. L., Chiew, F., Hughes, L., Newton, P. C. D., Schuster, S. S., … Whetton, P. (2014). Australasia. In V. R. Barros, C. B. Field, D. J. Dokken, M. D. Mastrandrea, K. J. Mach, T. E. Bilir, … L. L. White (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability. Part B: Regional aspects. Contribution of working group II to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 1371–1438). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Reser, J. P., Morrissey, S. A., & Ellul, M. (2011). The threat of climate change: Psychological response, adaptation, and impacts. In I. Weissbecker & A. Marsella (Eds.), *Climate change and human well-being* (pp. 19–42). New York, NY: Springer.

Reserve Bank of New Zealand. (2015). Household debt. Retrieved 15 April 2015, from http://www.rbnz.govt.nz/statistics/key_graphs/household_debt/

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*(1), 145–154. http://doi.org/10.1007/s11336-008-9102-z

Rindskopf, D. M. (2013). Testing 'small,' not null, hypotheses: Classical and Bayesian approaches. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 319–332). New York, NY: Psychology Press.

Ripley, B. (2014). MASS: Support functions and datasets for Venables and Ripley's MASS (Version 7.3-33). Retrieved from http://cran.r-project.org/web/packages/MASS/index.html

Robbins, M. C., Dewalt, B. R., & Pelto, P. J. (1972). Climate and behavior: A biocultural study. *Journal of Cross-Cultural Psychology*, *3*(4), 331–344. http://doi.org/10.1177/002202217200300402

Rock, D. J., & Hallmayer, J. F. (2008). The seasonal risk for deliberate self-harm. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, *29*(4), 191–201. http://doi.org/10.1027/0227-5910.29.4.191

Roe, R. A. (2008). Time in applied psychology: The study of 'what happens' rather than 'what is'. *European Psychologist*, *13*(1), 37–52. http://doi.org/10.1027/1016-9040.13.1.37

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. http://doi.org/10.1093/biomet/70.1.41

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York, NY: Russell Sage Foundation.

Rotton, J. (1986). Determinism redux: Climate and cultural correlates of violence. *Environment and Behavior*, *18*(3), 346–368. http://doi.org/10.1177/0013916586183003

Rotton, J., & Cohn, E. G. (2000). Violence is a curvilinear function of temperature in Dallas: A replication. *Journal of Personality and Social Psychology*, *78*(6), 1074–1081. http://doi.org/10.1037/0022-3514.78.6.1074

Rotton, J., & Cohn, E. G. (2003). Global warming and U.S. crime rates: An application of routine activity theory. *Environment and Behavior*, *35*(6), 802–825. http://doi.org/10.1177/0013916503255565

Rotton, J., & Cohn, E. G. (2004). Outdoor temperature, climate control, and criminal assault: The spatial and temporal ecology of violence. *Environment and Behavior*, *36*(2), 276–306. http://doi.org/10.1177/0013916503259515

Rubin, D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association*, *100*(469), 322–331. http://doi.org/10.1198/016214504000001880

Rubonis, A. V., & Bickman, L. (1991). Psychological impairment in the wake of disaster: The disaster–psychopathology relationship. *Psychological Bulletin*, *109*(3), 384–399. http://doi.org/10.1037/0033-2909.109.3.384

Ruuhela, R., Hiltunen, L., Venäläinen, A., Pirinen, P., & Partonen, T. (2008). Climate impact on suicide rates in Finland from 1971 to 2003. *International Journal of Biometeorology*, *53*(2), 167–175. http://doi.org/10.1007/s00484-008-0200-5

Salehyan, I. (2008). From climate change to conflict? No consensus yet. *Journal of Peace Research*, *45*(3), 315–326. http://doi.org/10.1177/0022343308088812

Salmond, C. E., & Crampton, P. (2012). Development of New Zealand's deprivation index (NZDep) and its uptake as a national policy tool. *Canadian Journal of Public Health*, *103*(8), S7–S11.

Salmond, C. E., Crampton, P., Sutton, F., & Atkinson, J. (2014). Socioeconomic Deprivation Indexes: NZDep and NZiDep. Retrieved 14 April 2015, from http://www.otago.ac.nz/wellington/departments/publichealth/research/hirp/otago020194.html

Schmidt, F. L., & Hunter, J. E. (1999). Theory testing and measurement error. *Intelligence*, *27*(3), 183–198. http://doi.org/10.1016/S0160-2896(99)00024-0

Schönbrodt, F. (2012). Visually weighted regression in R (à la Solomon Hsiang). Retrieved 5 February 2015, from http://www.nicebread.de/visually-weighted-regression-in-r-a-la-solomon-hsiang/

Senese, F. (2010). How can I calculate the energy required to cause a temperature rise? Retrieved from http://antoine.frostburg.edu/chem/senese/101/thermo/faq/energy-required-for-temperature-rise.shtml

Sherry, D. (2011). Thermoscopes, thermometers, and the foundations of measurement. *Studies In History and Philosophy of Science Part A*, *42*(4), 509–524. http://doi.org/10.1016/j.shpsa.2011.07.001

Sherwood, S. C., & Huber, M. (2010). An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences*, *107*(21), 9552–9555. http://doi.org/10.1073/pnas.0913352107

Sijtsma, K. (2008). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107–120. http://doi.org/10.1007/s11336-008-9101-0

Singer, E. (2006). Nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*(5), 637–645. http://doi.org/10.1093/poq/nfl034

Skinner, R. (1982). *Relativity for scientists and engineers*. Mineola, NY: Dover.

Skrondal, A., & Rabe-Hesketh, S. (2009). Prediction in multilevel generalized linear models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(3), 659–687. http://doi.org/10.1111/j.1467-985X.2009.00587.x

Sloman, S. A., & Lagnado, D. A. (2005). The problem of induction. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 95–116). New York, NY: Cambridge University Press.

Smith, K. R., Woodward, A., Campbell-Lendrum, D., Chadee, D. D., Honda, Y., Liu, Q., … Sauerborn, R. (2014). Human health: Impacts, adaptation, and co-benefits. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, … L. L. White (Eds.), *Climate change 2014: Impacts,* *adaptation, and vulnerability. Part A: Global and sectoral aspects. Contribution of working group II to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 709–754). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Smith, S. J., Edmonds, J., Hartin, C. A., Mundra, A., & Calvin, K. (2015). Near-term acceleration in the rate of temperature change. *Nature Climate Change*, *5*(4), 333–336. http://doi.org/10.1038/nclimate2552

Sohn, D. (1999). Experimental effects: Are they constant or variable across individuals? *Theory & Psychology*, *9*(5), 625–638. http://doi.org/10.1177/0959354399095003

Souêtre, E., Wehr, T. A., Douillet, P., & Darcourt, G. (1990). Influence of environmental factors on suicidal behavior. *Psychiatry Research*, *32*(3), 253–263. http://doi.org/10.1016/0165-1781(90)90030-9

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. http://doi.org/10.2307/1412159

Spence, A., Pidgeon, N., & Uzzell, D. (2009). Climate change: Psychology's contribution. *The Psychologist*, *22*(2), 108–111.

Statistics New Zealand. (2012a). Calendar year offence statistics: Calendar year population estimates. Retrieved 8 October 2013, from http://www.stats.govt.nz/tools_and_services/nzdotstat/recorded-crime-statistics/~/media/Statistics/tools/table-builder/crime-tables/SNZ%20Calendar%20Populations%2004Mar13.xls

Statistics New Zealand. (2012b). Estimated resident population for territorial authority areas, at 30 June (1996+) (annual-Jun). Retrieved 6 November 2012, from http://www.stats.govt.nz/infoshare/

Statistics New Zealand. (2013). Subnational population estimates tables. Retrieved 30 May 2014, from http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/subnational-pop-estimates-tables.aspx

Statistics New Zealand. (2014). 2013 Census QuickStats about culture and identity. Retrieved 29 January 2015, from http://www.stats.govt.nz/Census/2013-census/profile-and-summary-reports/quickstats-culture-identity/ethnic-groups-NZ.aspx

Statistics New Zealand. (n.d.-a). Estimated resident population (ERP), subnational population by ethnic group, age, and sex, at 30 June 1996, 2001, 2006, and 2013. Retrieved 14 January 2015, from http://nzdotstat.stats.govt.nz/wbos/Index.aspx?DataSetCode=TABLECODE7512

Statistics New Zealand. (n.d.-b). Infoshare. Retrieved 6 November 2012, from http://www.stats.govt.nz/infoshare/

Stern, N. (2007). *The economics of climate change: The Stern review*. Cambridge, United Kingdom: Cambridge University Press.

Stern, P. C. (2011). Contributions of psychology to limiting climate change. *American Psychologist*, *66*(4), 303–314. http://doi.org/10.1037/a0023235

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*(2684), 677–680. http://doi.org/10.1126/science.103.2684.677

Stoll-Kleemann, S., O'Riordan, T., & Jaeger, C. C. (2001). The psychology of denial concerning climate mitigation measures: Evidence from Swiss focus groups. *Global Environmental Change*, *11*(2), 107–117. http://doi.org/10.1016/S0959-3780(00)00061-3

Swim, J. K., Stern, P. C., Doherty, T. J., Clayton, S., Reser, J. P., Weber, E. U., … Howard, G. S. (2011). Psychology's contributions to understanding and addressing global climate change. *American Psychologist*, *66*(4), 241–250. http://doi.org/10.1037/a0023220

Tait, A., Henderson, R., Turner, R., & Zheng, X. (2006). Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. *International Journal of Climatology*, *26*(14), 2097–2115. http://doi.org/10.1002/joc.1350

Törő, K., Dunay, G., Bartholy, J., Pongrácz, R., Kis, Z., & Keller, É. (2009). Relationship between suicidal cases and meteorological conditions. *Journal of Forensic and Legal Medicine*, *16*(5), 277–279. http://doi.org/10.1016/j.jflm.2008.12.015

Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory & Psychology*, *19*(5), 579–599. http://doi.org/10.1177/0959354309341926

Uitenbroek, D. G. (1996). Seasonal variation in alcohol use. *Journal of Studies on Alcohol and Drugs*, *57*(1), 47–52. http://doi.org/10.15288/jsa.1996.57.47

Vaneckova, P., Hart, M., Beggs, P., & de Dear, R. (2008). Synoptic analysis of heat-related mortality in Sydney, Australia, 1993–2001. *International Journal of Biometeorology*, *52*(6), 439–451. http://doi.org/10.1007/s00484-007-0138-z van Gelder, T., & Port, R. F. (1995). It's about time: An overview of the dynamical approach to cognition. In R. F. Port & T. van Gelder (Eds.), *Mind as motion* (pp. 1–43). Cambridge, MA: MIT Press.

Ver Hoef, J. M., & Boveng, P. L. (2007). Quasi-poisson vs. negative binomial regression: How should we model overdispersed count data? *Ecology*, *88*(11), 2766–2772. http://doi.org/10.1890/07-0043.1

Vrij, A., Steen, J. V. D., & Koppelaar, L. (1994). Aggression of police officers as a function of temperature: An experiment with the fire arms training system. *Journal of Community & Applied Social Psychology*, *4*(5), 365–370. http://doi.org/10.1002/casp.2450040505

Vyssoki, B., Praschak-Rieder, N., Sonneck, G., Blüml, V., Willeit, M., Kasper, S., & Kapusta, N. D. (2012). Effects of sunshine on suicide rates. *Comprehensive Psychiatry*, *53*(5), 535–539. http://doi.org/10.1016/j.comppsych.2011.06.003

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. http://doi.org/10.3758/BF03194105

Wang, W., Rothschild, D., Goel, S., & Gelman, A. (in press). Forecasting elections with non-representative polls. *International Journal of Forecasting*. http://doi.org/10.1016/j.ijforecast.2014.06.001

Wang, X.-F. (2010). fANCOVA: Nonparametric analysis of covariance (Version 0.5–1). Retrieved from http://cran.r-project.org/web/packages/fANCOVA/index.html

Wansbeek, T., & Meijer, E. (2003). Measurement error and latent variables. In B. H. Baltagi (Ed.), *A companion to theoretical econometrics* (pp. 162–179). Oxford, United Kingdom: Blackwell.

Webb, A. E., Leighton, L. R., Schellenberg, S. A., Landau, E. A., & Thomas, E. (2009). Impact of the Paleocene-Eocene thermal maximum on deep-ocean microbenthic community structure: Using rank-abundance curves to quantify paleoecological response. *Geology*, *37*(9), 783–786. http://doi.org/10.1130/G30074A.1

Weinberg, C. R. (1993). Toward a clearer definition of confounding. *American Journal of Epidemiology*, *137*(1), 1–8.

Weisberg, S. (2005). *Applied linear regression*. Hoboken, NJ: Wiley.

Weitzman, M. L. (2009). On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics*, *91*(1), 1–19. http://doi.org/10.1162/rest.91.1.1

Whitmarsh, L. (2009). Behavioural responses to climate change: Asymmetry of intentions and impacts. *Journal of Environmental Psychology*, *29*(1), 13–23. http://doi.org/10.1016/j.jenvp.2008.05.003

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*(8), 594–604. http://doi.org/10.1037/0003-066X.54.8.594

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research & Evaluation*, *18*(11). Retrieved from http://www.pareonline.net/getvn.asp?v=18&n=11

Williams, M. N., Hill, S. R., & Spicer, J. (2013). In response to 'Temperature and violent crime in Dallas, Texas: Relationships and implications of climate change'. *Western Journal of Emergency Medicine*, *14*(5), 567–568. http://doi.org/10.5811/westjem.2013.5.17833

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Will climate change increase or decrease suicide rates? The differing effects of geographical, seasonal, and irregular variation in temperature on suicide incidence. *Climatic Change*, *130*(4), 519–528. http://doi.org/10.1007/s10584-015-1371-9

Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach*. Mason, OH: Cengage.

World Health Organisation. (2002). *The world health report 2002: Reducing risks, promoting healthy life*. Retrieved from http://www.who.int/whr/2002/en/

Wratt, D., Basher, R., Mullan, B., & Renwick, J. (n.d.). El Niño and climate forecasting. Retrieved 15 January 2015, from https://www.niwa.co.nz/our-science/climate/information-and-resources/clivar/elnino

Yang, A. C., Tsai, S.-J., & Huang, N. E. (2011). Decomposing the association of completed suicide with air pollution, weather, and unemployment data at different time scales. *Journal of Affective Disorders*, *129*(1–3), 275–281. http://doi.org/10.1016/j.jad.2010.08.010

Yang, B., Lester, D., & Yang, C.-H. (1992). Sociological and economic theories of suicide: A comparison of the U.S.A. and Taiwan. *Social Science & Medicine*, *34*(3), 333–334. http://doi.org/10.1016/0277-9536(92)90275-U

Yan, Y. Y. (2000). Geophysical variables and behavior: LXXXXIX. The influence of weather on suicide in Hong Kong. *Perceptual and Motor Skills*, *91*(2), 571–577. http://doi.org/10.2466/pms.2000.91.2.571

Yip, P. S. F., Chao, A., & Ho, T. P. (1998). A re-examination of seasonal variation in suicides in Australia and New Zealand. *Journal of Affective Disorders*, *47*(1–3), 141–150. http://doi.org/10.1016/S0165-0327(97)00135-3

Zinbarg, R. E. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ω_h. *Applied Psychological Measurement*, *30*, 121–144. http://doi.org/10.1177/0146621605278814

Zung, W. W. K., & Green, R. L. (1974). Seasonal variation of suicide and depression. *Archives of General Psychiatry*, *30*(1), 89–91. http://doi.org/10.1001/archpsyc.1974.01760070067010

9 Appendix A: Supplementary Materials for Study One (Assaults)

This Appendix comprises supplementary information about the method and results of Study One. This supplementary materials document was submitted along with manuscript itself for peer review. The version shown is that provided to the journal at the proofing stage, with minor formatting changes for consistency with the rest of the thesis, and some typographical errors corrected. References cited in these supplementary materials are included within the consolidated reference list at the end of the main body of this thesis. These supplementary materials are available from Springer via the citation below:

Williams, M. N., Hill, S. R., & Spicer, J. (2015). The relationship between temperature and assault in New Zealand [Supplemental materials]. *Climatic Change, 132*(4), 559–573. http://doi.org/10.1007/s10584-015-1438-7

9.1 Introduction

The following electronic supplementary materials provide additional technical information about the study's methods and results that, due to space restrictions, is not included in the main text of the article. Some substantive analyses that were not reported in the main text are also included here (see especially subsections 9.3.5, analysing the effect of controlling for national changes in household debt levels, 9.3.6, analysing the effect of heat waves, and 9.3.10, analysing the interaction between temperature and population density).

9.2 Methods

9.2.1 Recorded assault data.

The definition of assault in New Zealand was not explicitly given in the main article text. In New Zealand, "assault" can refer either to the threat of an application of force or to the actual application of force itself (which is defined separately as "battery" in some other jurisdictions). Specifically, assault is defined in New Zealand under section 2 (interpretation) of the Crimes Act (1961, p. 22) as follows:

> the act of intentionally applying or attempting to apply force to the person of another, directly or indirectly, or threatening by any act or gesture to apply such force to the person of another, if the person making the threat has, or causes the other to believe on reasonable grounds that he has, present ability to effect his purpose

The assaults examined in this study were those falling under part 8 (Crimes against the person) of the Crimes Act 1961. Indecent assaults, which fall under part 7 (Crimes against religion, morality and public welfare) were excluded. Assaults occurring in the Chatham Islands were excluded. The Chathams (population approximately 600) nominally fall within the Wellington police district, but are located around 680km southeast of mainland New Zealand.

9.2.2 Assault hospitalisation data.

In the main text we mention that the assault hospitalisation data analysed excludes emergency department stays of less than two days. This was due to inconsistencies in terms of how such stays were recorded by hospital staff. We based this exclusion criterion on a report by the Ministry of Health, focused on suicide and self-harm hospitalisations rather than assaults, which identified this problem in New Zealand hospitalisation data for the time period studied. Specifically, the Ministry of Health Suicide Facts report (2012) excludes “patients who were discharged from an emergency department after a length of stay of less than two days” (p. 65) in their analysis of self-harm hospitalisations. This same exclusion criterion was applied in our study following consultation with the Ministry of Health. Specifically, a hospitalisation was excluded from analysis when:

1. The date of discharge was only one day or less after the date of admission; *and*:
2. The health specialty code for the hospitalisation was one of the following: M05 (Emergency Medicine), M06 (Allied health / community emergency medicine), M07 (Generalist emergency medicine), or M08 (Specialist Intensive Care).

One additional issue relating to the assault hospitalisation data that is not discussed in the main article text is the fact that, over the course of the study period, there were some relatively important changes to the local government structure in New Zealand. These changes were the amalgamation of Christchurch City and Banks Peninsula in 2006 and the creation of the Auckland “supercity” in 2010. The districts recorded in the hospitalisations database corresponded to the districts (also known as territorial local authorities, or territorial authority areas) existing prior to these amalgamations. In order to use a district classification corresponding to the current districts of New Zealand, the seven previous districts corresponding to the current Auckland supercity were amalgamated into one Auckland district in the analyses performed. These districts were Auckland City, Manukau City, Waitakere City, North Shore City, Rodney, Franklin, and Papakura. Christchurch City and Banks Peninsula were similarly amalgamated.

One point of complication with respect to the amalgamation of districts was that the Franklin district was actually split between the Auckland, Waikato and Hauraki districts during the creation of the Auckland supercity. However, since the actual street addresses of patients in the hospitalisation database were not available, it was impossible to determine whether patients listed with domiciles in the Franklin district in fact resided in an area of the district later to become part of Auckland, part of Waikato, or part of Hauraki. All patients listed as living in Franklin were therefore classified as living in the (amalgamated) Auckland district.

9.2.3 Population data for use in analyses of recorded assaults.

Although the effect of population size was not of major interest in and of itself in the present study, it was important to have population estimates for use as statistical controls in the substantive analyses completed. These estimates needed to take the form of daily population estimates, given that this was the unit of temporal aggregation used in some of the analyses undertaken (e.g., analyses of the effects of irregular variation in temperature).

With respect particularly to the analyses of recorded assaults, the police districts of New Zealand do not correspond exactly to regional council or territorial authority (district) areas, making it somewhat challenging to acquire appropriate population estimates for the regions studied. Population estimates by police district were available from Statistics New Zealand (2012a), but not for the entire studied period, and only on an annual basis. The following steps were therefore undertaken to provide daily population estimates for the full studied period.

1. Resident population estimates as at 30 June 1996–2010 by police district were obtained from Statistics New Zealand (2012a), covering the majority of the period for which recorded crime data were available.
2. Population estimates for Counties Manukau, Auckland City, and Waitematā were combined into population estimates for an overarching Auckland region
3. National resident population estimates as at 30 June 1994–1996 were obtained from the Statistics New Zealand Infoshare database (Statistics New Zealand, n.d.-b)
4. The populations in the Auckland, Wellington and Canterbury regions for 1994 and 1995 were estimated by taking the proportion of the NZ population falling within each of these areas in 1996, and using these proportions to estimate the populations of these regions in 1994 and 1995 (using the estimated national resident population for these two years).
5. As mentioned above, the Wellington Police District nominally includes the Chathams Islands (approximately 680km off the coast), but assaults occurring in the Chathams were not included in substantive data analyses of the Wellington region. The population of the Chathams were therefore removed from the Wellington population estimates. Estimates of the resident population in the Chathams from 1996 to 2012 were obtained from the Statistics New Zealand Infoshare database. Chatham Islands population estimates for 1994 and 1995 were obtained by assuming that the population remained constant (at 760 people) from 1994 to 1996.
6. Linear interpolation was used to convert the annual resident population estimates by region into daily population estimates for the period 1 July 1994 to 31 July 2009.

9.2.4 Population data for analyses of assault hospitalisations.

Population estimates by district were required for analyses of assault hospitalisations. Unfortunately, population estimates by district were required for a longer period than that readily available in any single data source using consistent boundaries, and again only annual estimates were readily available. These problems again necessitated several steps of processing in order to produce daily population estimates for the period of interest. The following steps were used to obtain daily population estimates by district for the full study period (1993 to 2009).

1. Annual resident population estimates by district (using 1995 boundaries) as at 30 June 1995 to 2000 were obtained from Statistics New Zealand (R. Speirs, personal communication).
2. The estimates for Auckland City, Manukau City, Waitakere City, North Shore City, Rodney, Franklin, and Papakura were amalgamated into a single Auckland district. Similarly, the Christchurch City and Banks Peninsula estimates were combined to reflect the amalgamation of these two districts in 2006.
3. Annual resident population estimates by district (using 2013 boundaries) for 2001 to 2009 were obtained from the Statistics New Zealand Infoshare database.
4. The population estimates by district (with 2013 boundaries) were adjusted to apply to the 1995 boundaries. The 1995 boundaries were used (bar the district amalgamations noted above) as these boundaries were likely to be the most representative of those used when districts of domicile were actually recorded in the national minimum dataset for hospital events. The adjustment was accomplished by calculating the ratio of the 1995-boundary estimate to the 2013-boundary estimate for each district in 2000 (the latest year for which estimates were available for both sets of boundaries). The 2001–2010 population estimates were multiplied by this ratio to adjust for the very minor boundary differences occurring.
5. Finally, linear interpolation was used to convert the annual population estimates to daily estimates.

9.2.5 Meteorological data.

As mentioned in the article text, a single virtual climate station was selected to represent each region (for analyses of recorded assaults) or district (for analyses of assaults resulting in hospitalisation). This was accomplished by choosing the virtual station closest to the town centre of the most populous urban area or town within that district. Three steps were involved in the selection of climate stations.

Firstly, the largest town or urban area falling within each district or region was identified by obtaining estimated resident populations by urban area in New Zealand for 1996 to 2007 from Statistics New Zealand's Infoshare website. Each district or region was linked with the most populous urban area within it. In the cases of three districts (Hurunui, Kapiti Coast, and Tasman), manual selection of the largest town was necessary due either to no urban area falling within the district being listed in the Infoshare data, or the largest urban area not corresponding with the largest distinct town in the district (e.g., due to an "urban area" encompassing two towns in different districts).

Secondly, the centre of the largest town or urban area within each district or region was defined as per Google Maps, and its geographical co-ordinates obtained from the iTouchMaps tool (iTouchMap.com, 2013). The "city centres" thus identified were the cultural and commercial centres of each town or urban area, as opposed to representing geographical centroids.

Finally, the virtual climate station within the district or region that was nearest to the town centre was identified using the National Institute of Water and Atmospheric Research's CliFlo database (NIWA, n.d.-a). In a very small number of cases, the virtual climate station nearest the town centre fell outside the boundaries of the district itself. In these cases the station nearest to the town centre but within the district was used. The climate stations used to represent each district and region are listed in Table 4, along with their agent numbers within the CliFlo database.

Table 4

Virtual Climate Stations Utilised

		Virtual climate station		
Region/District	Most populous urban area	Agent no.	Latitude	Longitude
Regions				
Auckland	Auckland Urban Area	25396	-36.875	174.775
Wellington	Wellington Urban Area	28602	-41.275	174.775
Canterbury	Christchurch Urban Area	20810	-43.525	172.625
Districts (territorial authority areas)				
Ashburton District	Ashburton Urban Area	19146	-43.925	171.725
Auckland	Auckland Urban Area	25396	-36.875	174.775
Buller District	Westport Urban Area	18772	-41.775	171.625
Carterton District	Carterton Urban Area	30879	-41.025	175.525
Central Hawke's Bay District	Waipukurau Urban Area	31069	-39.975	176.575
Central Otago District	Alexandra Urban Area	12937	-45.225	169.375
Christchurch City	Christchurch Urban Area	20810	-43.525	172.625
Clutha District	Balclutha Urban Area	13350	-46.225	169.725
Dunedin City	Dunedin Urban Area	19446	-45.875	170.475
Far North District	Kaitaia Urban Area	20661	-35.125	173.275
Gisborne District	Gisborne Urban Area	30645	-38.675	177.975
Gore District	Gore Urban Area	13152	-46.075	168.925
Grey District	Greymouth Urban Area	19694	-42.475	171.225
Hamilton City	Hamilton Urban Area	30829	-37.775	175.275
Hastings District	Hastings Urban Zone	29002	-39.625	176.825
Hauraki District	Waihi Urban Area	29897	-37.375	175.825
Horowhenua District	Levin Urban Area	30825	-40.625	175.275
Hurunui District	Amberley	21366	-43.175	172.725
Hutt City	Lower Hutt Urban Zone	30748	-41.225	174.925
Invercargill City	Invercargill Urban Area	7643	-46.425	168.375
Kaikoura District	Kaikoura Urban Area	28055	-42.375	173.675
Kaipara District	Dargaville Urban Area	28571	-35.925	173.875
Kapiti Coast District	Paraparaumu	30219	-40.925	175.025
Kawerau District	Kawerau Urban Area	30029	-38.075	176.725
Mackenzie District	Twizel Community Urban Area	13690	-44.275	170.125
Manawatu District	Feilding Urban Area	30341	-40.225	175.575
Marlborough District	Blenheim Urban Area	27021	-41.525	173.975
Masterton District	Masterton Urban Area	28285	-40.975	175.675
Matamata-Piako District	Morrinsville Urban Area	30887	-37.675	175.525
Napier City	Napier Urban Zone	27434	-39.475	176.875
Nelson City	Nelson Urban Area	20719	-41.275	173.275
New Plymouth District	New Plymouth Urban Area	21442	-39.075	174.075
Opotiki District	Opotiki Urban Area	30066	-38.025	177.275
Otorohanga District	Otorohanga Urban Area	29745	-38.175	175.225

Palmerston North City	Palmerston North Urban Area	28276	-40.375	175.625
Porirua City	Porirua Urban Zone	27590	-41.125	174.825
Queenstown-Lakes District	Queenstown Urban Area	14372	-45.025	168.675
Rangitikei District	Marton Urban Area	27156	-40.075	175.375
Rotorua District	Rotorua Urban Area	27868	-38.125	176.225
Ruapehu District	Taumarunui Urban Area	28702	-38.875	175.275
Selwyn District	Rolleston Urban Area	20052	-43.575	172.375
South Taranaki District	Hawera Urban Area	21610	-39.575	174.275
South Waikato District	Tokoroa Urban Area	30961	-38.225	175.875
South Wairarapa District	Featherston Urban Area	28201	-41.125	175.325
Southland District	Winton Urban Area	10729	-46.125	168.325
Stratford District	Stratford Urban Area	21605	-39.325	174.275
Tararua District	Dannevirke Urban Area	27324	-40.225	176.125
Tasman District	Richmond	20430	-41.375	173.175
Taupo District	Taupo Urban Area	30999	-38.675	176.075
Tauranga City	Tauranga Urban Area	29942	-37.675	176.125
Thames-Coromandel District	Thames Urban Area	28786	-37.125	175.575
Timaru District	Timaru Urban Area	19840	-44.375	171.225
Upper Hutt City	Upper Hutt Urban Zone	30228	-41.125	175.075
Waikato District	Huntly Urban Area	30253	-37.575	175.175
Waimakariri District	Rangiora Urban Area	19946	-43.325	172.575
Waimate District	Waimate Urban Area	19832	-44.725	171.025
Waipa District	Cambridge Urban Zone	28244	-37.875	175.475
Wairoa District	Wairoa Urban Area	31126	-39.025	177.425
Waitaki District	Oamaru Urban Area	19617	-45.075	170.975
Waitomo District	Te Kuiti Urban Area	27114	-38.325	175.175
Wanganui District	Wanganui Urban Area	28141	-39.925	175.025
Wellington City	Wellington Urban Area	28602	-41.275	174.775
Western Bay of Plenty District	Te Puke Community Urban Area	29452	-37.775	176.325
Westland District	Hokitika Urban Area	19484	-42.725	170.975
Whakatane District	Whakatane Urban Area	27972	-37.975	176.975
Whangarei District	Whangarei Urban Area	21619	-35.725	174.325

9.2.6 Deprivation data.

As mentioned in the main text, New Zealand deprivation index (NZDep) data were obtained from the University of Otago (Salmond et al., 2014; see also Salmond & Crampton, 2012). Deprivation level was used as a control variable for a supplementary analysis concerned with the effect of geographical variation in temperature on the incidence of assaults resulting in hospitalisation. The NZDep is a measure of the socioeconomic deprivation of a geographical area, based on nine items from the New Zealand census (e.g., the number of people in the area aged 18–64 receiving a means-tested benefit; the number of people not living in their own home; the number of people without access to a car; etc.). The NZDep is reported in two forms: A numeric score scaled to have a mean of 1000 and standard deviation of 100 across meshblocks (the form used in our study), and an ordinal score ranging from 1 to 10.

NZDep scores were obtained in the forms of scores by census meshblock (the smallest geographical unit utilised by Statistics NZ when collecting data). Scores were obtained for each of the census years 1996, 2001, and 2006. NZDep scores by district for each of these three years were then obtained by averaging deprivation scores across the meshblocks falling within each district, while weighting for the size of the population of each meshblock. A single deprivation score for each district was then obtained by averaging deprivation scores across the three years of 1996, 2001, and 2006.

9.3 Results

The following supplementary materials for the results section mainly take the form of full listings of coefficients for the models reported in the article text, and in some cases the provision of the results of additional analyses not reported in the main text. However, we will start with some general additional information about the form and estimation of the models utilised.

9.3.1 General form of the models utilised.

The majority of statistical models reported in the article text took a similar form: They were generalised linear mixed models, with a Poisson distribution for the response variable, a log link function, and an estimated error term. The log link function means that the *log* of the expected assault count was assumed to be a linear (and additive) function of the parameters. The log link function is conventionally used when estimating Poisson models and related count-data models. Its use ensures that the predictions of the models are always non-negative, as must be the case for a count variable. The use of a freely estimated error term avoided any problems with overdispersion (see Coxe et al., 2009), a situation that occurs when the conditional variance of the response variable is greater than the conditional mean (the Poisson model assumes that the conditional mean and variance are equal).

As a consequence of the use of the log link function, the models used assume that expected assault counts are a *multiplicative* function of the predictors. This was a useful property, because it meant that the coefficients estimated were provided on a readily interpretable scale, and could easily be compared across analyses with different levels of aggregation. A coefficient, $\hat{\beta}$, from a model with a log link, can be translated into the proportional increase in the response variable as a result of a unit change in the predictor:

Proportional increase = $e^{\hat{\beta}} - 1$

When the value $\hat{\beta}$ is reasonably close to zero, the proportional increase in the response variable (e.g., assaults) for a unit change in the predictor (e.g., a 1°C increase in temperature) is approximately $\hat{\beta}$. For example, in the model regressing recorded assaults on irregular daily variation in temperature, the coefficient for temperature was 0.015. When translated to a proportional increase via the formula $e^{0.015} - 1$, the expected increase in assaults for a 1°C increase in temperature is simply 0.015 (or 1.5%).

Another consequence of the use of a log link function was that while the log of the expected assault count was assumed to be a linear function of the parameters, the relationship assumed between the predictors (e.g., temperature) and expected assault counts was not itself linear but, rather, exponential. This said, an exponential relationship between a predictor and a response variable will take a form very similar to a straight line as long as the coefficient for the predictor is reasonably small, and the range of values examined for the predictor is reasonably constrained. That was the case in this study, where the majority of models estimated implied a relationship between temperature and assaults that was very close to linear within the range of temperatures observed. As such, while it is certainly debatable whether the true relationship between temperature and assault incidence is better approximated as linear and additive or multiplicative and exponential, within the range of temperatures observed and coefficients estimated here, the distinction is not of particularly great importance.

However, many of the analyses used incorporated population size as a control variable, for which it seems particularly reasonable to specifically assume an additive rather than multiplicative effect. In other words, one would expect an increase in population of a given size to result in roughly the same number of extra assaults regardless of the population level prior to the increase. Therefore, a logarithmic transformation was applied to the population size predictor in all analyses reported.

9.3.2 Estimation of models.

As mentioned in the article text, the generalised linear mixed models reported were estimated using penalised quasi-likelihood. Bolker et al. (2009) suggests that this is an appropriate estimation method for generalised linear mixed models with a Poisson response, provided that the mean of the response variable for each combination of values on the predictors is greater than five. That was plausibly the case for all the analyses reported bar that pertaining to the relationship between irregular daily variation in temperature and assaults resulting in hospitalisation. However, for that model, estimation using the Laplace approximation (as suggested by Bolker et. al for counts with mean less than five) resulted in virtually identical results, justifying the use of penalised quasi-likelihood estimation. The alternative analysis was completed using the lme4 package in R (Bates et al., 2013).

9.3.3 Seasonal variation in temperature and recorded assaults.

As mentioned in the article, the effect of seasonal variation in temperature was estimated using a generalised linear mixed model estimated over the three regions (Auckland, Wellington, and Canterbury). An autoregressive term was used to deal with some moderate autocorrelation in the residuals. The addition of autoregressive terms (whether one term or more than one, e.g., AR7 in a model not reported here) resulted in little substantive change to the estimates, bar slightly increasing the standard errors.

In the article text, two models are reported: A model with just temperature and population size as predictors, and another also including a daylength control. Coefficients for these models are reported in Table 5. However, one issue of relevance to the seasonal analyses not mentioned in the main article text was a slight excess of assaults on public holidays in general (23% increase), and especially, on New Year's Day (112% increase). Since public holidays in New Zealand, including New Year's Day, fall disproportionately in warmer months, this represented a potential confound. The apparent excess of assaults occurring on New Year's Day may have occurred at least partly due to assaults being recorded with "partial dates", i.e., where the year of an assault was known at the time it was recorded, but not its actual date. A model including controls for public holidays and New Year's Day is therefore reported in Table 5, but resulted in little further substantive change to the estimated effect of temperature. The public holiday variable was the proportion of times that each day of the year was a public holiday over the study period (the date of some holidays varying from year to year). Where a public holiday was "Mondayised", both the actual date and the observed date were counted as holidays.

Table 5

Seasonal Variation in Temperature and Recorded Assaults: Model Coefficients

	Without daylength or holiday controls			With daylength control			With public holiday & New Year's day controls		
		95% CI			95% CI			95% CI	
	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper
Fixed effects									
Intercept	89.208	63.771	114.646	-10.080	-15.194	-4.966	-9.671	-14.822	-4.519
Log (population)	-6.276*	-8.153	-4.400	1.070	0.690	1.449	1.046	0.663	1.428
Tmean (°C)	0.014	0.011	0.017	-0.004	-0.008	$3.3\text{x}10^{-4}$	-0.004	-0.008	$9.8\text{x}10^{-5}$
Log (daylength)	-	-	-	0.461	0.378	0.543	0.424	0.345	0.503
New Year's Day	-	-	-	-	-	-	0.659	0.578	0.740
National public holiday	-	-	-	-	-	-	0.068	0.021	0.115
Random effects (SDs)									
Intercept | Region	3.465	1.853	6.481	0.159	0.072	0.352	0.161	0.071	0.365
Residual	1.819	1.739	1.903	1.751	1.675	1.830	1.546	1.476	1.618
Error correlations									
AR1 | Region	0.251	0.186	0.313	0.201	0.135	0.267	0.283	0.221	0.341

Notes. Response variable = number of assaults on each of the 365 days of the calendar year within each region (Auckland, Canterbury, and Wellington), summed over the entire study period (15 years). N = 365 days*3 regions = 1095. Poisson generalised linear mixed model with log link used. *The estimated effect of population for the model without daylength or holiday controls of -6.276 fails a "sanity check", implying that increases in population are associated with *decreases* in the number of assaults. However, constraining the effect of population size to 1 via an offset resulted in essentially no change to the coefficient for temperature.

9.3.4 Irregular daily variation in temperature and recorded assaults.

In the article text, estimates from four models estimating the effect of irregular daily variation in temperature on recorded assaults are mentioned: A model with month, weekday, population size and temperatures as predictors, a model also including relative humidity, a model in which the effect of temperature was permitted to vary randomly across regions, and a model including lagged effects of temperature. A full list of the coefficients for these models is provided in Table 6, with the exception of the coefficients for the model in which temperature was allowed to vary randomly (for brevity, since the article text already lists the random effect of temperature within each region).

Table 6

Irregular Daily Variation in Temperature and Recorded Assaults: Model Coefficients

	Without humidity or lagged effects			With humidity			With lagged temperature effects		
		95% CI			95% CI			95% CI	
	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper
Fixed effects									
Intercept	-12.471	-15.864	-9.077	-12.373	-15.750	-8.996	-12.501	-15.897	-9.106
Log (population)	1.148	0.895	1.401	1.141	0.889	1.393	1.150	0.897	1.404
Monday	-0.544	-0.562	-0.527	-0.544	-0.562	-0.526	-0.544	-0.562	-0.526
Tuesday	-0.490	-0.508	-0.472	-0.490	-0.508	-0.472	-0.489	-0.507	-0.472
Wednesday	-0.459	-0.477	-0.441	-0.459	-0.477	-0.442	-0.459	-0.477	-0.441
Thursday	-0.368	-0.385	-0.350	-0.368	-0.385	-0.351	-0.368	-0.385	-0.351
Friday	-0.203	-0.219	-0.186	-0.203	-0.219	-0.187	-0.203	-0.219	-0.186
Saturday	0.047	0.032	0.062	0.047	0.032	0.062	0.047	0.032	0.062
February	0.063	0.039	0.087	0.063	0.039	0.087	0.063	0.039	0.087
March	0.039	0.015	0.063	0.039	0.015	0.063	0.039	0.015	0.063
April	-0.108	-0.133	-0.083	-0.108	-0.133	-0.083	-0.108	-0.133	-0.083
May	-0.091	-0.115	-0.066	-0.091	-0.115	-0.066	-0.091	-0.115	-0.066
June	-0.152	-0.177	-0.127	-0.152	-0.177	-0.126	-0.152	-0.177	-0.127
July	-0.058	-0.082	-0.034	-0.058	-0.082	-0.034	-0.060	-0.084	-0.035
August	-0.028	-0.052	-0.003	-0.027	-0.052	-0.003	-0.027	-0.052	-0.003
September	-0.007	-0.031	0.018	-0.007	-0.031	0.018	-0.006	-0.031	0.018
October	0.034	0.010	0.058	0.033	0.009	0.057	0.034	0.010	0.058
November	0.070	0.046	0.094	0.070	0.045	0.094	0.070	0.046	0.094
December	0.099	0.075	0.122	0.098	0.074	0.122	0.099	0.075	0.123
Tmean, anomaly (°C)	0.015	0.012	0.017	0.016	0.013	0.018	0.015	0.012	0.018
Relative humidity, anomaly (%)	-	-	-	-7.4×10^{-4}	-1.3×10^{-3}	-2.3×10^{-4}	-	-	-
Humidity anomaly*Tmean anomaly	-	-	-	2.0×10^{-4}	-1.4×10^{-5}	4.1×10^{-4}	-	-	-
Tmean anomaly, lag 1	-	-	-	-	-	-	0.000	-0.004	0.004
Tmean anomaly, lag 2	-	-	-	-	-	-	0.000	-0.004	0.003
Tmean anomaly, lag 3	-	-	-	-	-	-	0.000	-0.003	0.004
Tmean anomaly, lag 4	-	-	-	-	-	-	0.001	-0.002	0.005
Tmean anomaly, lag 5	-	-	-	-	-	-	-0.002	-0.006	0.002
Tmean anomaly, lag 6	-	-	-	-	-	-	-0.002	-0.005	0.002

Tmean anomaly, lag 7	-	-	-	-	-	-	0.002	-0.001	0.005
Random effects (SDs)									
Intercept | Region	0.156	0.070	0.347	0.154	0.069	0.346	0.156	0.068	0.357
Intercept | Year in Region	0.079	0.064	0.099	0.079	0.063	0.099	0.079	0.064	0.099
Residual	1.256	1.242	1.270	1.255	1.242	1.269	1.256	1.242	1.269
Error correlations									
AR1 | Region	0.040	0.026	0.055	0.040	0.026	0.055	0.040	0.026	0.055

Notes. Response variable = number of assaults on a given date within each region (Auckland, Canterbury, and Wellington). *N* = 5510 dates*3 regions = 16,530. Poisson generalised linear mixed model with log link used.

9.3.5 Effect of household debt.

A reviewer of our manuscript suggested that household debt might be an important socioeconomic control variable to apply in temporal analyses. We did not include this variable as a control in our main analyses; rather, the random intercept across years within each district provided a flexible control for time-trending confounds. We did, however, complete a supplementary analysis assessing the effect of controlling for household debt, focusing again on possibly the most important analysis reported in our study (that of the effect of irregular daily variation in temperature on the incidence of recorded assaults). Specifically, at the suggestion of the reviewer, we obtained quarterly national observations of the percentage of nominal disposable income required to service household debt (from Reserve Bank of New Zealand, 2015) for use as a control variable. Over the study period, the household debt service percentage ranged quite widely, from 6.8% in 1994 to 14.3% in 2008. We used linear interpolation to convert the quarterly debt servicing percentages to daily observations, and entered the servicing percentage as a predictor into a mixed model. The model estimated was in other respects the same as the model estimating the effect of irregular daily variation in temperature on recorded assaults (without humidity or lagged effect controls), as reported above in Table 6. When the servicing percentage was added to this model, it had a positive relationship with assault incidence, $\hat{\beta}$ = 0.023, 95% CI [0.012, 0.034]. However, the estimated effect of temperature remained virtually unchanged, $\hat{\beta}$ = 0.015, 95% CI [0.012, 0.017].

9.3.6 Effect of heat waves.

Again at the suggestion of a reviewer, we also briefly examined the effect of heat waves on the incidence of recorded assaults. In New Zealand, temperatures rarely reach very high levels: The daily mean temperature never exceeded 28°C within the dataset used in this study. As such, New Zealand is not the ideal location for a study of the effect of heat waves, hence the inclusion of this analysis in the electronic supplementary materials rather than the main article. It is nevertheless possible to define a heat wave in a relative sense as a period of abnormally high temperatures for the regions studied: Specifically, we defined a heat wave as a period of five or more consecutive days in which the daily maximum temperature exceeded the average maximum temperature over the study period for that area by 5°C or more (see Frich et al., 2002). The average maximum temperature at the city centre of each region was 18.9°C in Auckland, 16.3°C in Wellington, and 17.0°C in Canterbury, resulting in heat wave thresholds of 23.9°C, 21.3°C, and 22.0°C respectively.

Thus defined, there were a total of 27 heat waves in Auckland, 28 in Canterbury, and 24 in Wellington, with a total of 547 heat wave days in the study period of 1 July 1994 to 31 July 2009. A mixed model was used to estimate the effect of heat waves, while controlling for the effects of month and weekday, with a random intercept across regions and an AR1 error structure (as in Table 6). In this model, the estimated effect of heat waves was $\hat{\beta} = 0.051$, 95% CI [0.023, 0.079], suggesting that heat waves resulted in an increase in assaults of more than 5%. However, when the effect of mean temperature was also controlled for, the 95% confidence interval for the effect of heat wave days included zero, $\hat{\beta} = 0.020$, 95% CI [-0.009, 0.048]. As such, there did not seem to be strong evidence that the prolonged high temperatures found in heat waves resulted in an increase in assaults that exceeded what would be expected given the previously established temperature effect. Model coefficients for the heat wave analyses are displayed in Table 7.

Table 7

Heat Waves and Recorded Assaults: Model Coefficients

	Model without temperature anomaly included			Model with temperature anomaly included		
		95% CI			95% CI	
	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper
Fixed effects						
Intercept	-12.516	-15.879	-9.152	-12.470	-15.865	-9.074
Log (population)	1.151	0.900	1.402	1.148	0.894	1.401
Monday	-0.545	-0.563	-0.528	-0.544	-0.562	-0.527
Tuesday	-0.491	-0.509	-0.473	-0.490	-0.508	-0.472
Wednesday	-0.460	-0.478	-0.442	-0.459	-0.477	-0.442
Thursday	-0.368	-0.385	-0.351	-0.368	-0.385	-0.350
Friday	-0.203	-0.220	-0.187	-0.203	-0.219	-0.187
Saturday	0.046	0.031	0.062	0.047	0.032	0.062
February	0.060	0.036	0.084	0.062	0.038	0.086
March	0.043	0.019	0.067	0.041	0.017	0.065
April	-0.101	-0.126	-0.075	-0.105	-0.130	-0.080
May	-0.084	-0.109	-0.058	-0.088	-0.113	-0.063
June	-0.144	-0.170	-0.119	-0.149	-0.174	-0.123
July	-0.051	-0.076	-0.026	-0.055	-0.080	-0.031
August	-0.020	-0.045	0.005	-0.024	-0.049	0.001
September	0.001	-0.024	0.026	-0.003	-0.028	0.022
October	0.041	0.016	0.066	0.037	0.012	0.062
November	0.077	0.052	0.102	0.073	0.048	0.097
December	0.104	0.080	0.129	0.102	0.077	0.126
Heat wave	0.051	0.023	0.079	0.020	-0.009	0.048
Tmean, anomaly (°C)	-	-	-	0.014	0.012	0.017
Random effects (SDs)						
Intercept | Region	0.156	0.067	0.362	0.156	0.071	0.340
Intercept | Year in Region	0.078	0.062	0.098	0.080	0.064	0.099
Residual	1.261	1.247	1.275	1.256	1.242	1.269
Error correlations						
AR1 | Region	0.047	0.030	0.063	0.040	0.026	0.055

Notes. Response variable = number of assaults on a given date within each region (Auckland, Canterbury, and Wellington). N = 5510 dates*3 regions = 16,530. Poisson generalised linear mixed model with log link used.

9.3.7 Recorded assaults and ENSO.

In the article text, the estimated effect of ENSO variability (operationalised as the Southern Oscillation Index or SOI) on summed monthly assaults across Auckland, Canterbury and Wellington is reported based on a linear mixed model. This model included a control for population size and a random effect for year. A log link function with a Poisson distribution was again used, and the logarithm of the number of days in each month was offset to account for months having differing lengths. The coefficients for this model are displayed in Table 8.

Table 8

ENSO and Recorded Assaults: Model Coefficients

		95% CI	
	$\hat{\beta}$	Lower	Upper
Fixed effects			
Intercept	-13.025	-18.068	-7.981
Log (population)	1.252	0.878	1.625
SOI	0.002	0.001	0.004
Random effects (SDs)			
Intercept | Year	0.048	0.030	0.077
Residual	3.349	3.005	3.731

Notes. Response variable = number of assaults in a given month summed across three regions (Auckland, Canterbury, and Wellington). *N* = 181 months (July 1994 to July 2009). Poisson generalised linear mixed model with log link used. Log of number of days in each month offset

9.3.8 Seasonal variation in temperature and assaults resulting in hospitalisation.

We now move on to supplementary information relating to the analyses of assaults resulting in hospitalisation. As mentioned in the article text, assaults resulting in hospitalisation were most common in the warmer months, and least common in the colder months. Mean daily assaults and temperature by month (across all districts) were not provided in the article text for brevity, but are displayed here in Figure 10.

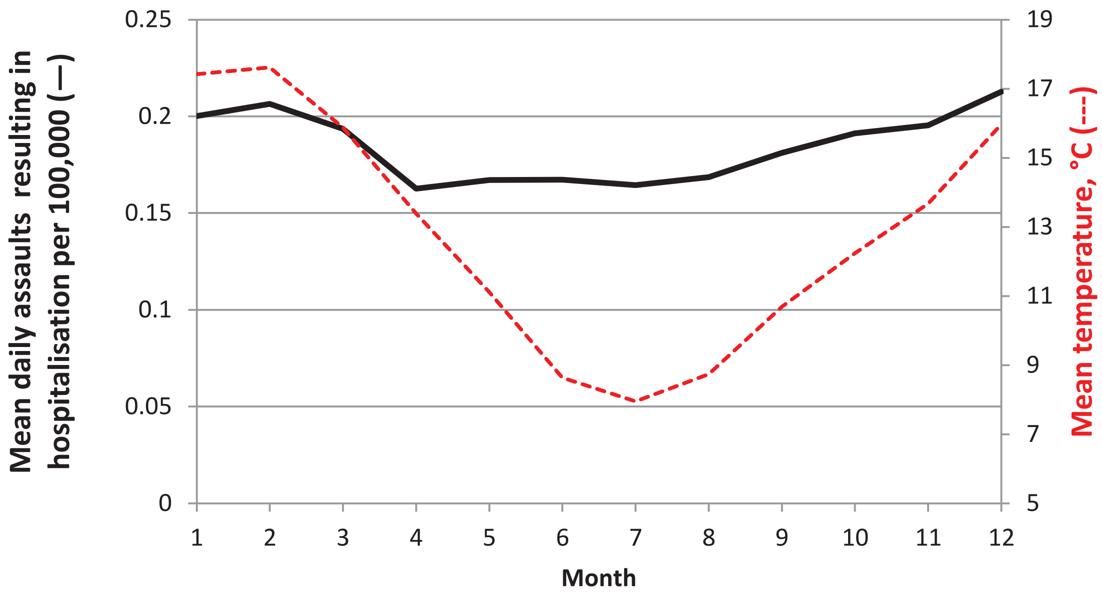

Figure 10. Mean daily assaults resulting in hospitalisation and temperatures by month.

The effect of seasonal variation in temperature on the incidence of assaults resulting in hospitalisation was again estimated using a generalised linear mixed model, estimated both with and without daylength controlled. The full list of coefficients for the models with and without daylength controlled is displayed in Table 9. A complication relating to the analysis of seasonal variation in assaults resulting in hospitalisation was that the original source of the data, the national minimum dataset for hospital events, allows partial dates to be entered for dates of injury. When an injury date is entered with a month and year but no day, it is recorded in the database as having occurred on the first day of the month. When a date is entered with no month or day, but with a year, it is recorded as having occurred on the first day of the year. This was a potential problem given the substantial excess of assaults that appeared to have happened on New Year's Day, usually a particularly warm day of the year in New Zealand (thus potentially introducing confounding, a problem also discussed previously with respect to the analyses of seasonal variation in recorded assaults). Indeed, in the data studied, the apparent number of assaults resulting in hospitalisation increased more than three-fold on New Year's Day. An alternative model was therefore estimated incorporating control variables for the first day of the year and the first day of the month (calendar controls), and is included in Table 9. A control for national public holidays was included in this model, since there are a number of public holidays in the warmest time of the year (Christmas/New Year). However, these controls did not substantially further alter the estimated effect of seasonal variation in temperature, which remained near zero.

Table 9

Seasonal Variation in Temperature and Assaults Resulting in Hospitalisation: Model Coefficients

	Without daylength controlled			With daylength controlled			With calendar controls		
		95% CI			95% CI			95% CI	
	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper
Fixed effects									
Intercept	-10.863	-11.949	-9.776	-11.755	-12.882	-10.628	-11.695	-12.823	-10.567
Log (population)	1.015	0.910	1.120	1.010	0.903	1.118	1.017	0.909	1.125
Tmean (°C)	0.023	0.020	0.026	0.004	-0.001	0.009	0.003	-0.002	0.008
Log (daylength)	-	-	-	0.475	0.375	0.576	0.420	0.322	0.519
Public holiday	-	-	-	-	-	-	0.101	0.028	0.173
First day of month	-	-	-	-	-	-	0.303	0.254	0.351
New Year's Day	-	-	-	-	-	-	1.009	0.887	1.130
Random effects (SDs)									
Intercept | District	0.429	0.359	0.511	0.439	0.370	0.521	0.440	0.369	0.526
Residual	1.053	1.043	1.063	1.049	1.040	1.059	1.022	1.013	1.031

Notes. Poisson generalised linear mixed model with log link. Response variable = number of assaults resulting in hospitalisation on each of the 365 days of the calendar year within each district, summed over the entire study period (17 years). N = 365 days*66 districts = 24,090.

9.3.9 Irregular daily variation in temperature and assaults resulting in hospitalisation.

The effect of irregular daily variation in temperature on assaults resulting in hospitalisation was estimated using generalised linear mixed models. Three models were mentioned in the article text: One with only temperature anomaly and population size as predictors, another adding lagged temperature effects, and finally a model examining the effect of relative humidity. The full set of coefficients from these models is displayed in Table 10.

Table 10

Irregular Daily Variation in Temperature and Assaults Resulting in Hospitalisation: Model Coefficients

	Without humidity or lagged effects			With lagged effects			With humidity		
		95% CI			95% CI			95% CI	
	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper
Fixed effects									
Intercept	-13.155	-14.209	-12.101	-13.171	-14.225	-12.117	-13.150	-14.201	-12.098
Log (population)	0.990	0.888	1.092	0.991	0.889	1.093	0.989	0.888	1.091
Tmean, anomaly (°C)	0.017	0.013	0.021	0.016	0.010	0.021	0.018	0.014	0.022
Tmean anomaly, lag 1	-	-	-	0.007	0.000	0.014	-	-	-
Tmean anomaly, lag 2	-	-	-	-0.006	-0.013	0.001	-	-	-
Tmean anomaly, lag 3	-	-	-	-0.002	-0.009	0.005	-	-	-
Tmean anomaly, lag 4	-	-	-	-0.001	-0.008	0.005	-	-	-
Tmean anomaly, lag 5	-	-	-	0.001	-0.006	0.008	-	-	-
Tmean anomaly, lag 6	-	-	-	-0.008	-0.015	-0.001	-	-	-
Tmean anomaly, lag 7	-	-	-	0.003	-0.002	0.009	-	-	-
Relative humidity, anomaly (%)	-	-	-	-	-	-	7.1×10^{-4}	1.7×10^{-3}	3.1×10^{-4}
Humidity anomaly* Tmean anomaly	-	-	-	-	-	-	6.0×10^{-4}	1.8×10^{-4}	1.0×10^{-3}
Random effects (SDs)									
Intercept | District	0.439	0.395	0.487	0.439	0.036	5.398	0.438	0.385	0.498
Intercept | Year in District	0.203	0.186	0.222	0.202	0.132	0.309	0.203	0.184	0.224
Residual	1.010	1.008	1.012	1.010	1.003	1.017	1.010	1.008	1.012

Notes. Response variable = number of assaults on a given date within each district. *N* = 6209 dates*66 districts = 409,794. Poisson generalised linear mixed model with log link.

9.3.10 Irregular daily variation in temperature, assaults resulting in hospitalisation, and population density.

Based on a comment from a reviewer, we also briefly examined whether population density moderated the effect of temperature on assault incidence. The dataset relating irregular daily variation in temperature to assaults resulting in hospitalisation was chosen for this purpose. Using this dataset allowed us to examine how the apparent effect of temperature might change in relation to population density using a form of temperature variation less susceptible to confounding (irregular variation), while drawing information about assault from a large number of districts with widely varying population densities. Population density was defined as the mean population of each district over the study period divided by its land area using 2001 boundaries. Population density ranged from 0.6 population/km^2 in the Mackenzie District to 1289 population/km^2 in Hamilton City. When the main effect of population density and its interaction were added to the leftmost model shown in Table 10 (i.e., without humidity or lagged effects) there was no evidence of either a main effect of population density, $\hat{\beta}$ = -4.6 x 10^{-4}, 95% CI [-9.7 x 10^{-4}, 4.4 x 10^{-5}], nor of an interaction with temperature, $\hat{\beta}$ = 5.2 x 10^{-6}, 95% CI [-1.3 x 10^{-5}, 2.3 x 10^{-5}]. Including population density made no substantive difference to the estimated effect of temperature: $\hat{\beta}$ = 0.016, 95% CI [0.011, 0.021].

9.3.11 Geographical variation in temperature and assaults resulting in hospitalisation.

As mentioned in the article text, a generalised linear model with a quasi-Poisson distribution and log link was used when estimating the relationship between geographical variation in temperature and assault counts by district (as summed over the entire study period). Three models are mentioned in the article text: One with just population size and mean temperature as predictors, another incorporating controls for age and ethnicity, and a final model also controlling for geographical differences in deprivation level. A full list of the coefficients for these models is provided in Table 11.

Table 11

Geographical Variation in Temperature and Assaults Resulting in Hospitalisation: Model Coefficients

	Without age & ethnicity controls			With age & ethnicity controls			With deprivation control		
		95% CI			95% CI			95% CI	
	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper	$\hat{\beta}$	Lower	Upper
Intercept	-4.838	-5.565	-4.120	-3.377	-7.566	0.796	-6.931	-13.474	-0.373
Log (population)	0.901	0.842	0.960	1.048	0.943	1.154	1.051	0.947	1.156
Tmean (°C)	0.113	0.053	0.175	-0.029	-0.089	0.031	-0.026	-0.085	0.034
% European	-	-	-	-0.008	-0.025	0.009	-0.004	-0.022	0.013
% Māori	-	-	-	0.022	0.009	0.035	0.016	0.001	0.031
% aged 15–39	-	-	-	-0.015	-0.063	0.033	-0.012	-0.060	0.036
% aged 40–64	-	-	-	-0.028	-0.101	0.044	-0.019	-0.092	0.053
% aged 65+	-	-	-	0.030	-0.021	0.081	0.022	-0.031	0.074
Deprivation	-	-	-	-	-	-	0.003	-0.001	0.007

Notes. Quasi-Poisson generalised linear model. Response variable: Assault counts by district over the entire study period. N = 66 districts. Quasi-Poisson dispersion parameter = 68.7 for model without age and ethnicity controlled, 33.7 for model with age and ethnicity controlled, and 32.9 for model with deprivation controlled. Age and ethnicity control variables obtained by averaging the percentage falling within each demographic group in each district across the 1996, 2001 and 2006 censuses. Deprivation scores are measured on a scale with a mean of 1000 and standard deviation of 100.

9.3.12 Trends in assaults resulting in hospitalisation.

For brevity's sake, trends in the annual incidence of assaults resulting in hospitalisation were not reported in the main article text. Over the study period there was a slight decline in the recorded incidence of assaults resulting in hospitalisation (as usual, excluding short emergency department stays), from a peak of 75.9 per 100,000 in 1994 to just 62.5 in 2009. This trend can be seen in Figure 11. When the analysis was restricted to the districts forming the Auckland, Wellington, and Canterbury regions and the year range to 1995–2008 for comparison with the recorded assault series, there was still a very slight downward trend (-0.1 assaults resulting in hospitalisation per 100,000 per year). The increase in recorded assaults over this period in these regions was therefore not mirrored by an increase in assaults resulting in hospitalisation, suggesting that the apparent downward trend in assault hospitalisations might primarily reflect changes in clinical practices over time.

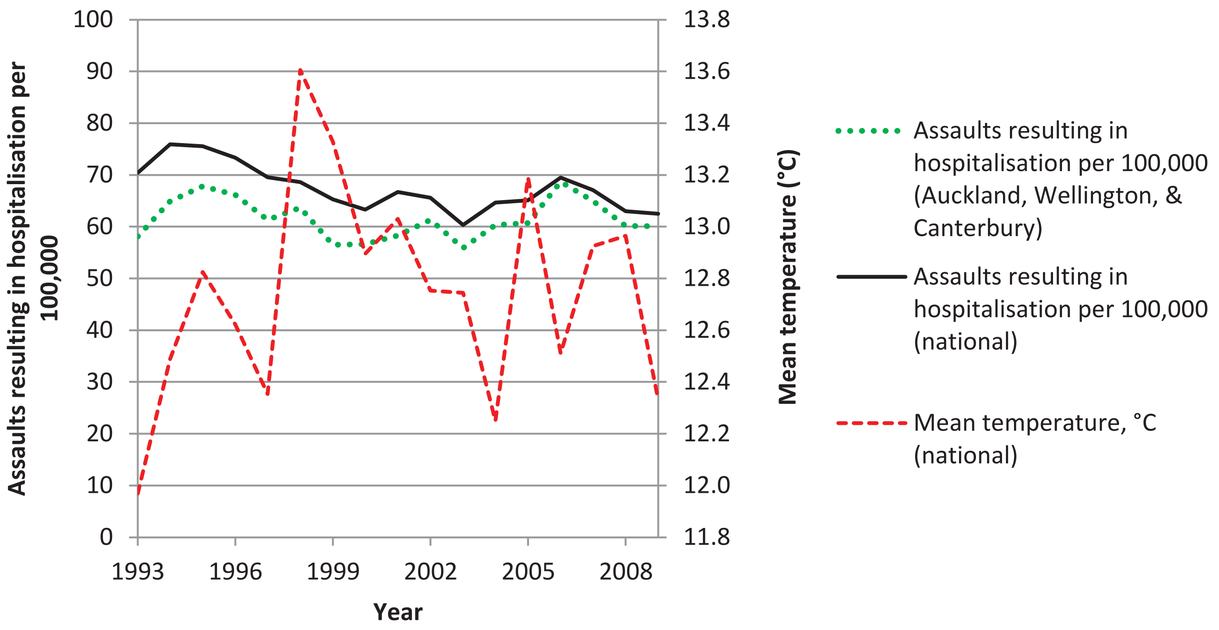

Figure 11. Trends in assaults resulting in hospitalisation over the study period.

10 Appendix B: Supplementary Materials for Study Two (Suicides)

The information found in this appendix was published electronically as a supplementary materials document along with Study Two. Once again, the supplementary materials provide additional information about the study's method and results that, due to space restrictions, could not be included in the main article text. The supplementary materials provided here are the same as those provided along with the final manuscript submitted for publication, except for some minor formatting changes for consistency with the rest of the thesis, and the correction of a small number of typographical errors. References cited in this appendix can be found in the consolidated reference list above. The final published version of these supplementary materials can be found at:

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Will climate change increase or decrease suicide rates? The differing effects of geographical, seasonal, and irregular variation in temperature on suicide incidence [Supplemental material]. *Climatic Change*, *130*(4), 519–528. http://doi.org/10.1007/s10584-015-1371-9

10.1 Additional Information about Methods

10.1.1 Population data.

The geographical boundaries used when determining population estimates were, for the most part, those existing as at 1995 (a time point near the middle of the suicide data series). Two exceptions to this rule were an amalgamation of Banks Peninsula with Christchurch City, and the amalgamation of seven districts into the Auckland "supercity". These amalgamations were conducted in order to better represent current geographical classifications in New Zealand. Maps showing the districts of New Zealand can be found on the Local Government New Zealand website (LGNZ, 2015). Annual population estimates for years prior to 1991 were estimates of resident population by district based on the national de facto population in these years, under the assumption that each district had the same share of the national population in these years as they did in 1991. Population estimates from 1991 to 2000 were estimates of resident population by district (1995 boundaries) obtained directly from Statistics New Zealand (R. Speirs, personal communication). Estimates for 1996 onwards were estimates by district, based initially on 2013 boundaries, and adjusted to apply to the 1995 boundaries. Annual population estimates were converted to daily population estimates using linear interpolation. When entered into statistical models as a predictor, population was log transformed and centred around the logarithm of the mean population across districts.

Ethnicity and age data were obtained from the Statistics New Zealand website (Statistics New Zealand, 2013). Specifically, the following variables were obtained for each of the 1996, 2001 and 2006 censuses:

- Number of residents aged 0–14
- Number aged 15–39
- Number aged 40–64
- Number aged 65+
- Number Māori
- Number Pacific peoples
- Number Middle Eastern/Latin American/African
- Number Asian
- Number European or Other ethnicity (including New Zealander)

Residents could fall into more than one ethnic category. The number falling into each age and ethnic category over the three censuses was then summed, and converted into a mean percentage for that category and district over the whole study period. For the purposes of use in statistical models, the percentage of residents aged 0–14 variable was excluded, as one of the age categories would be redundant when the other three were controlled. Similarly, only the two largest ethnic groups (European/other and Māori) were included as predictors, in order to keep the number of predictor variables reasonably constrained. The percentages falling into each age and ethnic category were log-transformed before entry into statistical models. In combination with the use of a log link function in the substantive analyses, this allowed the age and ethnicity variables to have additive rather than multiplicative effects.

10.1.2 Meteorological data.

Recordings from NIWA's virtual climate network (VCN) were used in preference to physical station measurements in order to ensure that a weather station was available to represent each district, and avoided missing data. The VCN provides estimates of a variety of weather variables on a regular 5km grid, using a thin-plate smoothing spline model for spatial interpolations (see NIWA, n.d.-b; Tait et al., 2006). The one exception to this rule was that physical weather station measurements were used for the offshore Chatham Islands, for which VCN measurements were not available. The only missing data present in the meteorological dataset was for the Chatham Islands (64 dates with missing temperature observations, and 116 with missing radiation observations). The VCN station closest to the town centre of the most populous town or city within each district was used to represent the given district.

10.1.2.1 Definitions of meteorological variables.

The mean temperature for each district, as used in the analysis of geographical variation in temperature, was simply calculated as the arithmetic mean of the daily mean temperatures observed in the district over the entire study period.

$$GeoTemp_j = \frac{1}{n}\sum_{m=1}^{n} Temperature_{mj}$$

Where:

$GeoTemp_j$ is the mean temperature for district *j*

$Temperature_{mj}$ is the observed mean temperature for district *j* and date *m*

n is the number of dates in the study period (7305)

When entered into statistical models, the geographical mean temperatures were also centred around the overall national mean temperature.

The operational definition of the seasonal component of variation in temperature is given verbally in the main article text, but is provided here in symbolic form for additional clarity:

$$SeasTemp_{ij} = \left[\frac{1}{20}\sum_{k=1988}^{2007}\left(Temperature_{ijk}\right)\right] - GeoTemp_j$$

Where:

$SeasTemp_{ij}$ is the seasonal component of temperature variation for day of the calendar year *i* (e.g., "December 23") and district *j*

$Temperature_{ijk}$ is the observed temperature on day *i* in year *k* in district *j*

$GeoTemp_j$ is the mean temperature for district *j* over the entire study period, as defined above.

February 29 was excluded in analyses of seasonal variation, given the small sample size representing this day of the calendar year (i.e., just five leap years in the study period).

The definition of the irregular component of temperature variation (i.e., temperature anomalies) is also given in the main text, but is provided below in symbolic form for additional clarity:

$$TempAnom_{ijk} = Temperature_{ijk} - SeasTemp_{ij} - GeoTemp_j$$

Where:

$TempAnom_{ijk}$ is the temperature anomaly for day *i* in year *k* in district *j*

$Temperature_{ijk}$ is the observed temperature on day *i* in district *j* in year *k*

$SeasTemp_{ij}$ is the seasonal component of temperature variation for day of the calendar year *i* and district *j*, as defined above

$GeoTemp_j$ is the mean temperature for district *j* over the entire study period, as defined above.

10.1.3 Data analysis.

As mentioned in the main text, Poisson generalised linear mixed models were used for the analyses of seasonal and irregular variation in temperature and suicides. The Poisson model was used because the response variable (suicide incidence) was a count variable, with small counts for most dates and districts. The Poisson model assumes that, for any given combination of values on the predictors, the mean and variance of the response variable will be equal. In reality, variance greater than the mean can often occur, an assumption breach known as overdispersion. Overdispersion was checked by calculating the ratio of the Pearson chi-square fit statistic to the residual degrees of freedom. In the absence of overdispersion this ratio should be close to one

(Coxe et al., 2009), which was indeed the case for all of the mixed models estimated. In these generalised linear mixed models, the only random effect specified was a random intercept across districts—i.e., allowing the mean suicide rate to vary randomly across districts.

For the analysis of geographical variation in temperature and suicide, no random intercept across districts was required, since the goal was to try and explain geographical differences rather than control for them. Furthermore, the response variable was simply the sum of suicides across the entire study period for each district. Therefore, a mixed model was not required, and a generalised linear model was instead estimated for the analysis of geographical variation. A Poisson model was not reported for this analysis, due to the presence of overdispersion if such a model was used (ratio of Pearson chi-square statistic to degrees of freedom = 2.7). Instead, a negative binomial model was used. The negative binomial model is a more flexible count-data model, in which the variance of the response variable is assumed to be a quadratic function of the mean, rather than strictly equal to the mean (see Hilbe, 2011). A quasi-Poisson model, an alternative count-data model in which the variance is assumed to be a multiplicative function of the mean, also produced very similar results.

When modelling time series, serial dependence of model errors can be a problem. Error dependence was checked for by examining the autocorrelations amongst the residuals from the fitted models. These autocorrelations were generally very small (e.g., well under 0.1 even at very short lags), and as such correlated error structures were not specified for the models reported.

10.2 Additional Information about Results

In the main article text, full listings of coefficients for the models estimated were not reported; instead, only the coefficients of most interest (e.g., the estimated effects of temperature) were reported. In this section, a more complete description of the estimated models is provided.

10.2.1 Irregular variation in temperature.

As mentioned in the article text, the estimated effect of irregular variation in temperature was nearly identical across a range of model specifications, including a control for radiation and use of two different statistical models (negative binomial and

Poisson). The model reported below in Table 12 is a Poisson generalised linear mixed model, and includes a control for radiation.

Table 12

Irregular Variation in Temperature and Suicides, Contemporaneous Effects: Coefficients for Poisson Generalised Linear Mixed Model

Predictor	Coefficient	95% confidence interval Lower	95% confidence interval Upper	*SD* (for random effects)
Random effects				
Intercept | District	-3.880	-3.920	-3.839	0.112
Fixed effects				
Log population (centred)	0.968	0.931	1.006	-
Radiation (MJ/m^2, centred)	0.002	-0.001	0.004	-
Temperature anomaly (°C)	0.018	0.009	0.027	-

Notes. The SD column indicates the standard deviation for those coefficients allowed to vary randomly across grouping units (e.g., the standard deviation of the intercept, which was allowed to vary across districts). *N* = 7305 dates * 67 districts - 149 days with missing temperature or radiation observations = 489,286.

A model including lagged effects of temperature (up to a 7-day lag) was also estimated, as mentioned in the article. Coefficients for this model are reported in Table 13. No radiation control was incorporated in this model, given the absence of any substantial lagged effects of temperature even in the absence of such a control. As is visible in the table, the contemporaneous effect of temperature remains positive despite the incorporation of lagged effects, but none of the lagged effects themselves are statistically significant.

Table 13

Irregular Variation in Temperature and Suicides, with Lagged Effects: Coefficients for Poisson Generalised Linear Mixed Model

		95% confidence interval		
Predictor	Coefficient	Lower	Upper	*SD* (for random effects)
Random effects				
Intercept | District	-3.880	-3.921	-3.839	0.112
Fixed effects				
Log population (centred)	0.968	0.930	1.006	-
Temperature anomaly	0.016	0.004	0.027	-
Temperature anomaly (lag 1)	0.009	-0.005	0.023	-
Temperature anomaly (lag 2)	-0.012	-0.026	0.002	-
Temperature anomaly (lag 3)	0.001	-0.012	0.015	-
Temperature anomaly (lag 4)	0.007	-0.007	0.021	-
Temperature anomaly (lag 5)	-0.002	-0.016	0.012	-
Temperature anomaly (lag 6)	0.002	-0.012	0.016	-
Temperature anomaly (lag 7)	-0.007	-0.018	0.005	-

Notes. *N* = 7305 dates * 67 districts - 64 days with missing temperature observations = 489,371.

10.2.2 Seasonal variation in temperature.

As mentioned in the article text, seasonal variation in suicides was captured as the total number of suicides occurring in each of the 365 days of the calendar year within each district over the entire 20-year study period. Seasonal variation in temperature was defined as explained above (section 10.1.2.1) and in the article text. Coefficients for the model relating seasonal variation in temperature and suicides are reported in Table 14.

Table 14

Seasonal Variation in Temperature and Suicides: Coefficients for Poisson Generalised Linear Mixed Model

Predictor	Coefficient	95% confidence interval Lower	95% confidence interval Upper	*SD* (for random effects)
Random effects				
Intercept | District	-0.877	-0.918	-0.837	0.111
Fixed effects				
Log population (centred)	0.986	0.948	1.024	-
Seasonal temperature (°C)	0.000	-0.006	0.007	-

Notes. *N* = 365 days of the calendar year * 67 districts = 24,455.

10.2.3 Geographical variation in temperature.

Three models relating geographical variation in temperature and suicide were reported in the article text: One with only population as a control variable, one adding age and ethnicity controls, and another model also including a radiation control. The model without such controls is reported in Table 15.

Table 15

Geographical Variation in Temperature and Suicides: Coefficients for Poisson Generalised Linear Model

Predictor	Coefficient	95% confidence interval Lower	95% confidence interval Upper
Intercept	5.028	4.987	5.069
Log population (centred)	0.984	0.945	1.024
Mean temperature, °C (centred)	0.005	-0.022	0.032

Notes. *N* = 67 districts.

The model incorporating age and ethnicity controls is reported in Table 16. In this model, the estimated effect of geographical variation in temperature becomes negative.

Table 16

Geographical Variation in Temperature and Suicides: Coefficients for Poisson Generalised Linear Model with Age and Ethnicity Controls

		95% confidence interval	
Predictor	Coefficient	Lower	Upper
Intercept	1.195	-6.184	8.573
Log population (centred)	1.002	0.948	1.056
Log of percent population aged 15–39	0.563	-0.387	1.511
Log of percent population aged 40–64	0.672	-0.524	1.866
Log of percent population aged 65+	0.533	0.179	0.888
Log of percent population European/Other	-0.469	-1.031	0.098
Log of percent population Māori	0.110	-0.017	0.237
Mean temperature, °C (centred)	-0.034	-0.067	-0.001

Notes. N = 67 districts.

A model also including a control for radiation is reported in Table 17.

Table 17

Geographical Variation in Temperature and Suicides: Coefficients for Poisson Generalised Linear Model with Age, Ethnicity and Radiation Controls

		95% confidence interval	
Predictor	Coefficient	Lower	Upper
Intercept	0.846	-6.622	8.312
Log population (centred)	0.997	0.942	1.053
Log of percent population aged 15–39	0.607	-0.354	1.566
Log of percent population aged 40–64	0.717	-0.489	1.923
Log of percent population aged 65+	0.551	0.192	0.909
Log of percent population European/Other	-0.473	-1.037	0.097
Log of percent population Māori	0.114	-0.014	0.242
Radiation, MJ/m^2 (centred)	-0.020	-0.073	0.033
Mean temperature, °C (centred)	-0.026	-0.065	0.012

Notes. N = 67 districts.

Finally, one of the reviewers for this article also suggested an examination of the relationship between geographical variation in temperature and suicide incidence, subdivided by season. This analysis was not reported in the article for brevity, but is reported below. The model reported in Table 16 was therefore estimated again within each of the four seasons (that is, with suicides summed and temperature averaged within each of the four seasons). The seasons in New Zealand are generally defined as follows: December–February is summer, March–May is autumn, June–August is winter, and September–November is spring. The estimated effect of geographical variation in temperature within each of the four seasons is reported in Table 18. As is visible in the table, there were some differences in the point estimate of the effect of temperature across the seasons, but the 95% confidence intervals for the effect of temperature across the four seasons all overlap substantially. As such, there is no strong evidence of differences in the effect of geographical variation in temperature by season.

Table 18

Estimated Effect of Geographical Variation in Temperature, by Season: Coefficients from Poisson Generalised Linear Mixed Model

	Estimated effect of temperature		
		95% confidence interval	
	Coefficient	Lower	Upper
Summer	-0.080	-0.129	-0.031
Spring	0.003	-0.059	0.066
Winter	-0.031	-0.078	0.016
Autumn	-0.042	-0.084	0.000

Notes. The control variables utilised were the same as in Table 16 (control variable coefficients not shown). A separate model was fit for each season. N = 67 districts.

11 Appendix C: Supplementary Materials for Study Three (Self-Harm)

The information found in this appendix was published electronically (after peer review) as a supplementary electronic materials document along with Study Three. The supplementary materials provided here are the same as those provided along with the final manuscript submitted for publication, except for some minor formatting changes for consistency with the rest of the thesis. The references cited in this appendix can be found in the consolidated reference list. The final published version of these supplementary materials can be found at:

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Do hotter temperatures increase the incidence of self-harm hospitalisations? [Supplemental material]. *Psychology, Health & Medicine*. Advance online publication. http://doi.org/10.1080/13548506.2015.1028945

11.1 Introduction

The following supplementary materials provide extra information about the method and results of the study that is not included in the main text of the article due to space restrictions.

11.2 Additional Information about Methods

11.2.1 Meteorological data.

As mentioned in the article text, a virtual climate station was selected for each district by choosing the virtual station closest to the town centre of the most populous urban area or town within that district. Specifically, this was accomplished by firstly obtaining the estimated resident population by urban area in New Zealand for 1996 to 2007 from Statistics New Zealand's Infoshare website (Statistics New Zealand, 2012b), and linking each district with the most populous urban area within it. In the cases of four districts (Chatham Islands, Hurunui, Kapiti Coast, and Tasman), manual selection of the largest town was necessary due either to no centre defined as an "urban area" being listed in the Infoshare data, or the largest urban area not corresponding with the largest distinct town in the district (e.g., due to an "urban area" encompassing two towns in different districts). Secondly, the centre of the largest town or urban area was defined as per Google Maps, and its geographical co-ordinates obtained from the iTouchMaps tool. Finally, the virtual climate station within the district nearest the town centre was identified using the NIWA CliFlo database (NIWA, n.d.-a). In a very small number of cases, the virtual climate station nearest the town centre fell outside the boundaries of the district itself; in these cases the station nearest to the town centre but within the district was used. The climate stations used to represent the climate of each district are listed in Table 19.

As mentioned in the text, there was one district for which virtual climate station data was not available: The Chatham Islands. This was due to the Chathams, which are approximately 680km southeast of the North Island, falling outside the range of the virtual climate network. A physical station was therefore utilised for the territory: The Chatham Islands automatic weather station (AWS). There was a small quantity of missing data for this station: 69 days with missing temperature measurements, and 116 with missing radiation measurements. Missing data was dealt with by listwise deletion given its small quantity.

11.2.2 Geographical units/districts.

The territorial authority areas ("districts") recorded in the hospitalisations database corresponded to the territorial authority areas existing in New Zealand prior to the creation of the Auckland "supercity" in 2010. In order to use a district classification corresponding to the current districts of New Zealand, the seven previous districts corresponding to the current Auckland supercity were amalgamated into one Auckland district in the analyses performed. These districts were Auckland City, Manukau City, Waitakere City, North Shore City, Rodney, Franklin, and Papakura.

One point of complication was that the Franklin district was actually split between the Auckland, Waikato and Hauraki districts during the creation of the Auckland supercity. However, since the actual street addresses of patients in the hospitalisation database were not available, it was impossible to determine whether patients listed with domiciles in the Franklin district in fact resided in an area of the district later to become part of Auckland, part of Waikato, or part of Hauraki. All patients listed as living in Franklin were therefore classified as living in the (amalgamated) Auckland district.

One other district amalgamation occurring toward the end of the study period was that of Banks Peninsula with Christchurch City in 2006. These two districts were also amalgamated for the purposes of the current study.

Table 19

Virtual Climate Stations Used

District	Biggest urban area	VCN[1] agent no.	Latitude	Long-itude	Km from town centre
Ashburton District	Ashburton Urban Area	19146	-43.925	171.725	2.6
Auckland	Auckland Urban Area	25396	-36.875	174.775	3.1
Buller District	Westport Urban Area	18772	-41.775	171.625	2.8
Carterton District	Carterton Urban Area	30879	-41.025	175.525	0.5
Central Hawke's Bay District	Waipukurau Urban Area	31069	-39.975	176.575	2.8
Central Otago District	Alexandra Urban Area	12937	-45.225	169.375	2.7
Chatham Islands Territory	Waitangi	6191[2]	-43.950	-176.567	0.5
Christchurch City	Christchurch Urban Area	20810	-43.525	172.625	1.2
Clutha District	Balclutha Urban Area	13350	-46.225	169.725	1.8
Dunedin City	Dunedin Urban Area	19446	-45.875	170.475	2.2
Far North District	Kaitaia Urban Area	20661	-35.125	173.275	1.1
Gisborne District	Gisborne Urban Area	30645	-38.675	177.975	4.0
Gore District	Gore Urban Area	13152	-46.075	168.925	3.1
Grey District	Greymouth Urban Area	19694	-42.475	171.225	3.0
Hamilton City	Hamilton Urban Area	30829	-37.775	175.275	1.4
Hastings District	Hastings Urban Zone	29002	-39.625	176.825	2.0
Hauraki District	Waihi Urban Area	29897	-37.375	175.825	2.3
Horowhenua District	Levin Urban Area	30825	-40.625	175.275	1.0
Hurunui District	Amberley	21366	-43.175	172.725	2.1
Hutt City	Lower Hutt Urban Zone	30748	-41.225	174.925	2.3
Invercargill City	Invercargill Urban Area	7643	-46.425	168.375	2.1
Kaikoura District	Kaikoura Urban Area	28055	-42.375	173.675	2.9
Kaipara District	Dargaville Urban Area	28571	-35.925	173.875	1.8
Kapiti Coast District	Paraparaumu	30219	-40.925	175.025	1.8
Kawerau District	Kawerau Urban Area	30029	-38.075	176.725	2.5
Mackenzie District	Twizel Community Urban Area	13690	-44.275	170.125	2.4
Manawatu District	Feilding Urban Area	30341	-40.225	175.575	0.6
Marlborough District	Blenheim Urban Area	27021	-41.525	173.975	1.7
Masterton District	Masterton Urban Area	28285	-40.975	175.675	3.0
Matamata-Piako District	Morrinsville Urban Area	30887	-37.675	175.525	2.0
Napier City	Napier Urban Zone	27434	-39.475	176.875	3.7

Nelson City	Nelson Urban Area	20719	-41.275	173.275	0.9
New Plymouth District	New Plymouth Urban Area	21442	-39.075	174.075	2.2
Opotiki District	Opotiki Urban Area	30066	-38.025	177.275	2.2
Otorohanga District	Otorohanga Urban Area	29745	-38.175	175.225	2.0
Palmerston North City	Palmerston North Urban Area	28276	-40.375	175.625	2.9
Porirua City	Porirua Urban Zone	27590	-41.125	174.825	1.6
Queenstown-Lakes District	Queenstown Urban Area	14372	-45.025	168.675	1.2
Rangitikei District	Marton Urban Area	27156	-40.075	175.375	0.7
Rotorua District	Rotorua Urban Area	27868	-38.125	176.225	2.5
Ruapehu District	Taumarunui Urban Area	28702	-38.875	175.275	1.5
Selwyn District	Rolleston Urban Area	20052	-43.575	172.375	1.8
South Taranaki District	Hawera Urban Area	21610	-39.575	174.275	1.5
South Waikato District	Tokoroa Urban Area	30961	-38.225	175.875	0.9
South Wairarapa District	Featherston Urban Area	28201	-41.125	175.325	0.9
Southland District	Winton Urban Area	10729	-46.125	168.325	2.1
Stratford District	Stratford Urban Area	21605	-39.325	174.275	1.5
Tararua District	Dannevirke Urban Area	27324	-40.225	176.125	3.0
Tasman District	Richmond	20430	-41.375	173.175	4.1
Taupo District	Taupo Urban Area	30999	-38.675	176.075	1.3
Tauranga City	Tauranga Urban Area	29942	-37.675	176.125	3.8
Thames-Coromandel District	Thames Urban Area	28786	-37.125	175.575	3.3
Timaru District	Timaru Urban Area	19840	-44.375	171.225	3.4
Upper Hutt City	Upper Hutt Urban Zone	30228	-41.125	175.075	0.4
Waikato District	Huntly Urban Area	30253	-37.575	175.175	2.4
Waimakariri District	Rangiora Urban Area	19946	-43.325	172.575	2.9
Waimate District	Waimate Urban Area	19832	-44.725	171.025	2.0
Waipa District	Cambridge Urban Zone	28244	-37.875	175.475	1.9
Wairoa District	Wairoa Urban Area	31126	-39.025	177.425	1.3
Waitaki District	Oamaru Urban Area	19617	-45.075	170.975	2.5
Waitomo District	Te Kuiti Urban Area	27114	-38.325	175.175	1.3
Wanganui District	Wanganui Urban Area	28141	-39.925	175.025	2.0
Wellington City	Wellington Urban Area	28602	-41.275	174.775	1.3
Western Bay of Plenty District	Te Puke Community Urban Area	29452	-37.775	176.325	1.2
Westland District	Hokitika Urban Area	19484	-42.725	170.975	1.1
Whakatane District	Whakatane Urban Area	27972	-37.975	176.975	2.8
Whangarei District	Whangarei Urban Area	21619	-35.725	174.325	0.1

Notes. [1]VCN = Virtual Climate Network. [2]Physical weather station utilised (Chatham Islands AWS).

Our usage of the term "district" is technically at odds with the way this term is used in local government in New Zealand. We use the term to refer to all territorial authority areas, while in local government parlance there is a distinction between "district councils" and "city councils", both being types of territorial local authority. We use “district” as a generic term to refer to all types of territorial authority areas for the sake of brevity.

11.2.3 Population estimates.

The fact that subnational population estimates were required for a longer period than that readily available in any single data source using consistent boundaries necessitated several adjustments. The following steps were used to obtain daily population estimates by district for the full study period (1993 to 2009).

1. Annual resident population estimates by district (using 1995 boundaries) as at 30 June 1995 to 2000 were obtained from Statistics New Zealand (R. Speirs, personal communication).
2. The estimates for Auckland City, Manukau City, Waitakere City, North Shore City, Rodney, Franklin, and Papakura were amalgamated into a single Auckland district. Similarly, the Christchurch City and Banks Peninsula estimates were combined to reflect the amalgamation of these two districts in 2006.
3. Annual resident population estimates by district (using 2013 boundaries) for 2001 to 2009 were obtained from the Statistics New Zealand Infoshare database.
4. The population estimates by district (with 2013 boundaries) were adjusted to apply to the 1995 boundaries. The 1995 boundaries were used (bar the district amalgamations noted above) as these boundaries were likely to be the most representative of those used when districts of domicile were actually recorded in the national minimum dataset for hospital events. The adjustment was accomplished by calculating the ratio of the 1995-boundary estimate to the 2013-boundary estimate for each district in 2000 (the latest year for which estimates were available for both sets of boundaries). The 2001–2010 population estimates were multiplied by this ratio to adjust for the very minor boundary differences occurring.
5. Finally, linear interpolation was used to convert the annual population estimates to daily estimates.

National population estimates were also required for the brief analysis of trends in self-harm and temperature reported at the end of the results section (section 4.4.4). National resident population estimates were obtained from the Statistics New Zealand Infoshare database for 1993 to 2009 (mean of year ending 31 December).

The size of the populations of the districts studied varied widely. For example, the mean population over the study period ranged from over 1.2 million in Auckland to just 707 in the Chatham Islands. Five districts had a mean population of over 100,000 (averaged over the study period), these districts being Hamilton City, Wellington City, Christchurch City, Dunedin City, and Auckland. Over the study period, the population of New Zealand increased from approximately 3.6 million in 1993 to 4.3 million in 2009.

11.3 Additional Information about Results

The following additional information about the results focuses primarily on issues relating to the distributional assumptions of the methods employed. In addition, the final subsection provides a re-analysis of the study data using Bayesian rather than frequentist methods.

11.3.1 Effects of geographical variation in temperature.

As mentioned in the article text, overdispersion was checked for all the models reported in the results section by calculating the ratio of the Pearson chi-square statistic to its degrees of freedom for each fitted model. This ratio should be close to one in the absence of overdispersion. In the case of the geographical comparison analysis, the use of a negative binomial model (as reported in the article text) was used in order to avoid overdispersion that was present when a simpler Poisson model was used. A Poisson model for counts of self-harm across districts implied a negative effect of temperature, $\hat{\beta}$ = -0.081, 95% CI [-0.087, -0.074], but the chi-square to degrees of freedom ratio was 52.9:1, indicating very substantial overdispersion (this ratio should be close to one). A negative binomial model was used instead, and reported in the article text.

The chi-square to degrees of freedom ratio of 1.09 for this model suggested that the negative binomial model was reasonably successful in combating overdispersion. However, a negative binomial model is not the only model that can be used in the presence of overdispersion. Another possibility is the quasi-Poisson model. Whereas the negative binomial model specifies the predicted variance as a quadratic function of the predicted mean of the response variable (for a given level of the predictor variables), the quasi-Poisson model specifies the conditional variance as a multiplicative function of the predicted mean. The quasi-Poisson model tends to give greater weight to observations with large counts of the response variable (Ver Hoef & Boveng, 2007). This means that a quasi-Poisson model would give greater weight to districts with larger populations and therefore more incidents of self-harm.

Because both the negative binomial model and the quasi-Poisson model are plausible choices for the analysis of geographical variation, we include coefficients from the estimated models with a quasi-Poisson distribution (see Table 20). The results shown can be compared with the coefficients when using the negative binomial distribution in Table 1 of the main text. In the uncontrolled model for geographical variation reported in the main text and using a negative binomial model, the estimated effect of temperature was negative, $\hat{\beta}$ = -0.004, but with a confidence interval spanning zero. With a quasi-Poisson distribution, the estimated effect of temperature in the uncontrolled model is much more strongly negative, $\hat{\beta}$ = -0.081, 95% CI [-0.128, -0.032]. However, in the controlled model, the point estimate of the effect of temperature is positive but with a confidence interval spanning zero in both the negative binomial and quasi-Poisson analyses, indicating greater consistency in substantive results across the two methods used.

Table 20

Coefficients for Quasi-Poisson Geographical Variation Models

	Uncontrolled model			Controlled model		
		95% CI			95% CI	
Coefficient	Est.	lower	upper	Est.	lower	upper
Intercept	6.652	6.577	6.726	6.599	6.516	6.679
Log population*	0.953	0.902	1.005	1.023	0.926	1.120
Temperature (°C)	-0.081	-0.128	-0.032	0.025	-0.052	0.102
Radiation (MJ/m^2)				0.010	-0.083	0.103
Percentage European				-0.018	-0.045	0.011
Percentage Māori				-0.008	-0.029	0.014
Percentage Asian				-0.091	-0.154	-0.027
Percentage aged 15–39				0.095	0.043	0.147
Percentage aged 40–64				0.011	-0.059	0.081
Percentage aged 65+				0.103	0.054	0.152

Notes. *Each variable was centred around its cross-district mean, except for population which was centred around the logarithm of the cross-district mean population estimate. Generalised linear model with quasi-Poisson distribution and log link used. N = 67 districts. Quasi-Poisson dispersion parameter = 52.9 for uncontrolled model, and 30.4 for controlled model.

11.3.2 Effect of seasonal variation in temperature.

11.3.2.1 Centring procedure.

As noted in the text, the method for capturing seasonal variation in temperature in a statistical model was to calculate a mean temperature for each of the 365 days of the calendar year for each district. This calculation does in fact include both geographical and seasonal variation in the resulting temperature estimates for each calendar day and district. However, the use of random intercepts across districts in the linear mixed model meant that in effect the influence of geographical variation in temperature—and all other geographical factors influencing self-harm rate—were controlled. Thus, the resulting estimate for the effect of temperature refers solely to seasonal variation in temperature. Effectively identical results were found if the seasonal norm temperatures were group-mean centred around the mean temperature within each district.

11.3.2.2 Partial dates.

There was an apparent excess of self-harm incidents on the first day of each year (mean 11.5 incidents per day), and, to a slight extent, the first days of the remaining months (mean 8.1 incidents per day). In comparison, there were 7.6 incidents per day in the remainder of the calendar year. This was likely due to the fact that the national minimum dataset allows partial dates to be entered for dates of injury. When an injury date is entered with a month and year but no day, it is recorded in the database as having occurred on the first day of the month. When a date is entered with no month or day, but with a year, it is recorded as having occurred on the first day of the year. This was a potential problem given the substantial excess of self-harm incidents that appeared to have happened on New Year's Day, a particularly warm day of the calendar year in New Zealand. However, adding a control for first day of year and first day of month to the models reported in the seasonal variation subsection resulted in no substantial change to the results shown (i.e., coefficients changed by less than 0.001).

11.3.3 Effects of irregular variation in temperature.

One potential point of concern with relation to distributional assumptions in the analysis of the effects of irregular daily variation in temperature was that the mean number of self-harm incidents per date and district was quite low (M = 0.11), with the majority (91%) of dates within each district having no self-harm incidents resulting in hospitalisation. The Poisson model is well suited to counts of rare events, however, and the χ^2/df ratio of 0.97 (in both the model without radiation controlled as well as the model with radiation controlled) indicated no evidence of violation of the assumptions of the Poisson model. Inspection of a plot of nationally averaged daily temperature anomalies versus daily self-harm rate, with a line of best fit estimated using a loess smoother, also showed no evidence of non-linearity in the relationship between temperature anomalies and self-harm rate.

11.3.4 Alternative Bayesian analysis.

The analyses in the main text are reported without any use of statistical significance testing, given the many known problems with such tests (J. Cohen, 1994; Gigerenzer, Krauss, & Vitouch, 2004; Gill, 1999; Wagenmakers, 2007). Instead, confidence intervals were used as the primary inferential tool. Nevertheless, the analyses used were frequentist in nature. In the frequentist interpretation of probability, probability refers to the limit of the relative frequency of some event over a large number of trials. Frequentist confidence intervals have a rather unintuitive interpretation: If we were to repeat a study a very large number of times, with a new dataset each time, and calculate a 95% confidence interval for a parameter in each case, then 95% of these intervals should include the true population parameter. We technically cannot say, however, that that there is a 95% probability that a *specific* calculated 95% confidence interval actually contains the true population parameter. The true parameter is a fixed quantity, unchanging over multiple trials, and thus the frequentist interpretation of probability does not allow us to make probability statements about whether or not it falls in a specific interval. Similarly, a frequentist interpretation of probability does not permit statements about the probability that some hypothesis is true or false.

Another interpretation of probability is the Bayesian interpretation, where probability refers to a degree of belief in some proposition. This broader interpretation of probability allows for the making of probability statements about whether a particular hypothesis is true, or about whether a parameter falls within some interval. The Bayesian interpretation of probability is linked to Bayes' theorem, which shows how *prior* (existing) knowledge or beliefs can be combined with new observed data to produce a *posterior* probability distribution. This posterior probability distribution may simply be a figure indicating the probability that a proposition is true, or it may be a continuous probability distribution. One might obtain, for example, a continuous posterior distribution for the effect of temperature on self-harm incidence, with this distribution indicating which values of the effect of temperature are more and less probable (taking into account both prior knowledge and the data observed). A general introduction to Bayesian data analysis can be found in Kruschke (2010b).

As an alternative to the analyses reported in the main text, Bayesian analyses were also performed. The main outputs reported from these analyses are *credible* intervals. Credible intervals are the Bayesian analogue of frequentist confidence intervals, but unlike confidence intervals they have a very intuitive interpretation: There is a 95% probability that the true parameter falls within the 95% credible interval (given the priors specified and data observed). This is one major advantage of Bayesian data analysis, although in some cases (e.g., when an uninformative[31] prior distribution is

[31] An uninformative prior distribution for a parameter is one that gives no or very limited information about which values of the parameter are more probable than others. It indicates a state of ignorance about the value of the parameter (prior to seeing the data at hand). For example, a (very) uninformative prior for a regression coefficient would be a uniform distribution on $[-\infty, \infty]$, indicating that the parameter can fall anywhere on the real number line, with all values being equally probable.

used), frequentist confidence intervals will approximate Bayesian credible intervals (see Greenland & Poole, 2013).

In the main text (section 4.4), frequentist models estimating the effects of geographical, seasonal, and irregular variation in temperature were reported. For each of the three components of variation, two models were reported: One with just population controlled, and one with additional control variables (including radiation). Bayesian versions of each of these models were also estimated, with their key outputs reported here, and exactly the same control variables and random effects specified in each case as in the main article. Bayesian generalised linear models were fit for the geographical analysis using the bayesglm function in the R package arm version 1.7–03 (Gelman et al., 2014). The quasi-Poisson distribution was used in the geographical analysis instead of the negative binomial, given that the negative binomial distribution is not as readily implemented in R without using additional software. Bayesian generalised linear mixed models were estimated using the package MCMCglmm version 2.21 (Hadfield, 2010).

One of the main challenges in a Bayesian analysis is the specification of appropriate prior distributions. The prior probability distributions specified for the models reported incorporated informative priors only for the effect of temperature (the output of most interest), and the effect of population (a particularly important control variable for which strong prior information was available). For the effect of temperature (regardless of whether geographical, seasonal, or irregular), a normal distribution was specified with mean zero and standard deviation of 0.01, indicating that an increase in self-harm incidence of between -1 and +1% for every 1°C was most plausible, and an effect of more than 3% in either direction was very implausible. In other words, the effect of temperature was expected to be reasonably small, given the findings previously reported in this area, and the relatively small effects of temperature on suicide found in a similar study in New Zealand (Williams et al., 2015). The prior for the effect of log population was a normal distribution with $M = 1$, and $SD = 0.03$, given that it seemed reasonable to expect the incidence of self-harm to be approximately proportional to population size. Uninformative priors ($M = 0$, variance = 1.0×10^{10}) were specified for the other fixed effects (the intercepts, and the demographic controls in the geographical analysis). In other words, essentially no pre-existing knowledge or beliefs about the likely values of these parameters were incorporated into the analysis. The default uninformative prior in the MCMCglmm package was also used for random effects, being an (improper) inverse Wishart with parameters nu = 0 and V = 1.

Coefficients from the Bayesian models are reported in Table 21. For brevity, we report only the point estimates and credible intervals for the effect of temperature from each model, and provide the frequentist confidence interval from the corresponding original analysis in the main text for comparison. For the most part, the results of the Bayesian and frequentist analyses are very similar: The point and interval estimates for the effect of seasonal and irregular daily variation in temperature are nearly identical, for example. In the geographical analysis, the point estimate and credible interval limits for the effect of temperature are shrunk toward zero in comparison to the frequentist analysis. This reflects the fact that the smaller sample size (N = 67 districts) in the geographical analysis meant that the prior distribution—and its assumption that small effects of temperature were most plausible—exerted a stronger effect on the results than was the case for the seasonal and irregular analyses.

Table 21

Comparing Bayesian and Frequentist Models: Coefficients for Temperature

	Frequentist model			Bayesian model		
		95% confidence interval			95% credible interval	
Model	Est.	Lower	Upper	Est.	Lower	Upper
Geographical variation in temperature						
Temp effect in uncontrolled model	-0.004	-0.062	0.054	-0.002	-0.008	0.005
Temp effect in controlled model	0.050	-0.035	0.135	2.5×10^{-4}	-0.006	0.007
Seasonal variation in temperature						
Temp effect in model without radiation	0.005	0.002	0.008	0.005	0.002	0.007
Temp effect in model with radiation	0.001	-0.004	0.005	0.001	-0.003	0.006
Irregular daily variation						
Temp effect in model without radiation	0.007	0.003	0.011	0.006	0.002	0.010
Temp effect in model with radiation	0.007	0.003	0.012	0.007	0.001	0.011

Notes. Bayesian generalised linear model (quasi-Poisson, log link) used for geographical variation models. Bayesian generalised linear mixed model (Poisson, log link) used for remaining models. The specification of control variables and random effects was the same as in the corresponding frequentist models reported in the main text.

12 Appendix D: Temperature and Violent Crime in Dallas, Texas—Published Commentary

Included below is a letter to the editor published in the *Western Journal of Emergency Medicine*. The letter was written by my supervisors and I during my PhD candidature and relates to the general topic of the thesis. However, it does not form a part of the overall planned narrative of the thesis. It is therefore included only as an appendix. In the letter, we critique a study by Gamble and Hess (2012) that reported an analysis of the relationship between temperature and violent crime in Dallas. Our letter reports a simple simulation study demonstrating some important implications of one of the models reported by Gamble and Hess. This simulation study was mentioned briefly earlier in section 6.5.2.4.

The full citation to the final version of this publication is included below. The *Western Journal of Emergency Medicine* is an open-access journal, meaning that permission was not required for the reproduction of the manuscript here. The version shown is the final accepted manuscript. The references cited are included in the consolidated reference list provided earlier in the thesis.

Williams, M. N., Hill, S. R., & Spicer, J. (2013). In response to ‘Temperature and violent crime in Dallas, Texas: Relationships and implications of climate change’. *Western Journal of Emergency Medicine*, *14*(5), 567–568. Retrieved from https://escholarship.org/uc/item/3sf1c7vp.pdf

To the editor:

We were interested to read Gamble and Hess’s (2012) study finding that the daily incidence of violent crime in Dallas increased with temperatures up to 90°F (32.2°C), but decreased above this threshold. On this basis, their abstract surprisingly concludes that “higher ambient temperatures expected with climate change…. are not likely to be accompanied by markedly higher rates of violent crime” (p .239). This conclusion contrasts with the findings of previous studies (Anderson et al., 2000; Anderson & DeLisi, 2011; Rotton & Cohn, 2003).

Unfortunately, the authors did not attempt to actually estimate the magnitude of future warming that would be sufficient to bring about a decrease in aggregate annual violent crime, which will differ from the inflection point of the relationship between daily temperature and violent crime. We therefore used the piecewise regression model reported by the authors in order to investigate how annual aggravated assault incidence in Dallas is likely to be affected by changes in mean temperature. We focus on aggravated assault given that this was the crime for which a marked effect of temperature was reported. Temperature data for Dallas International Airport in 1999 was collected from the NCDC (National Oceanic and Atmospheric Administration, n.d.). 1999 was used as a reference point being the last year in the series investigated by Gamble and Hess.

The simulation was conducted as follows. For each of a range of hypothetical annual temperature anomalies from -5 to +20°F, the annual anomaly was added to the actual temperature in each day of 1999 to obtain an annual series of daily temperatures. The piecewise regression model was then used to obtain the predicted number of assaults for each day of the series. These were then summed over the course of the year.

Our simulation suggests that the mean temperature in Dallas would have to increase by around 13°F (7.2°C) before subsequent temperature increases would begin to reduce annual aggravated assaults. At this point, the model predicts an extra 145 annual aggravated assaults per 100,000 population in comparison to a world with zero warming. Before this point, temperature increases would continue to increase assaults. Notably, a temperature increase of 13°F would be substantially greater than the warming likely by the end of the 21st century on the basis of regional climate projections for central North America (Christensen et al., 2007). As such, the inflection point in the temperature-violence relation appears to occur at too high a temperature to be of much comfort for those concerned with the implications of climate change for human violence in the medium term.

However, it is important to note that this analysis provides only conditional predictions about how many extra assaults are likely to arise in Dallas given a particular magnitude of warming, in comparison to an identical Dallas without this warming. The world of the future will be different from today’s in many ways other than simply being warmer. An unconditional forecast of future violent crime rates would need to take into account multiple predictors of crime, as well as temporal trends unrelated to global warming—such as the decreasing trend in violent crime in Dallas over the last two decades (Federal Bureau of Investigation, 2010).

13 Appendix E: Statements of Contribution for Publications

Massey University requires that a Statement of Contribution (form DRC 16) is included for each publication included in a doctoral thesis. A copy of this form is provided below for each of the four publications contained in this thesis. Information about the contributions of my supervisors and I to the publications included can also be found in the Preface at the beginning of the thesis.

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Matthew Neil Williams

Name/Title of Principal Supervisor: Dr. Stephen Hill

Name of Published Research Output and full reference:

Williams, M. N., Hill, S. R., & Spicer, J. (2015). The relationship between temperature and assault in New Zealand. Climatic Change, 132(4), 559–573. http://doi.org/10.1007/s10584-015-1438-7

In which Chapter is the Published Work: 2

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate:

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  Matt Williams designed the study, collected data, selected and conducted data analyses, and wrote the article.

Matt Williams Digitally signed by Matt Williams DN: cn=Matt Williams, o=Massey University, ou=School of Psychology, email=m.n.williams@massey.ac.nz, c=NZ Date: 2015.06.24 10:16:49 +12'00'

Candidate's Signature

24/6/2015

Date

Stephen Hill Digitally signed by Stephen Hill DN: cn=Stephen Hill, o=Massey University, ou=School of Psychology, email=s.r.hill@massey.ac.nz, c=NZ Date: 2015.06.24 15:46:55 +12'00'

Principal Supervisor's signature

24/6/2015

Date

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Matthew Neil Williams

Name/Title of Principal Supervisor: Dr. Stephen Hill

Name of Published Research Output and full reference:

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Will climate change increase or decrease suicide rates? The differing effects of geographical, seasonal, and irregular variation in temperature on suicide incidence. Climatic Change, 130(4), 519–528. http://doi.org/10.1007/s10584-015-1371-9

In which Chapter is the Published Work: 3

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate:

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  Matt Williams designed the study, collected data, selected and conducted data analyses, and wrote the article.

Matt Williams Digitally signed by Matt Williams
DN: cn=Matt Williams, o=Massey University, ou=School of Psychology, email=m.n.williams@massey.ac.nz, c=NZ
Date: 2015.06.24 10:19:22 +12'00'

Candidate's Signature

24/6/2015

Date

Stephen Hill Digitally signed by Stephen Hill
DN: cn=Stephen Hill, o=Massey University, ou=School of Psychology, email=s.r.hill@massey.ac.nz, c=NZ
Date: 2015.06.24 15:46:29 +12'00'

Principal Supervisor's signature

24/6/2015

Date

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Matthew Neil Williams

Name/Title of Principal Supervisor: Dr. Stephen Hill

Name of Published Research Output and full reference:

Williams, M. N., Hill, S. R., & Spicer, J. (2015). Do hotter temperatures increase the incidence of self-harm hospitalisations? Psychology, Health & Medicine. Advance online publication. http://doi.org/10.1080/13548506.2015.1028945

In which Chapter is the Published Work: 4

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate:

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  Matt Williams designed the study, collected data, selected and conducted data analyses, and wrote the article.

Matt Williams
Digitally signed by Matt Williams
DN: cn=Matt Williams, o=Massey University, ou=School of Psychology, email=m.n.williams@massey.ac.nz, c=NZ
Date: 2015.06.24 10:23:06 +12'00'

Candidate's Signature

24/6/2015

Date

Stephen Hill
Digitally signed by Stephen Hill
DN: cn=Stephen Hill, o=Massey University, ou=School of Psychology, email=s.r.hill@massey.ac.nz, c=NZ
Date: 2015.06.24 15:46:04 +12'00'

Principal Supervisor's signature

24/6/2015

Date

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Matthew Neil Williams

Name/Title of Principal Supervisor: Dr. Stephen Hill

Name of Published Research Output and full reference:

Williams, M. N., Hill, S. R., & Spicer, J. (2013). In response to 'Temperature and violent crime in Dallas, Texas: Relationships and implications of climate change'. Western Journal of Emergency Medicine, 14(5), 567–568. Retrieved from https://escholarship.org/uc/item/3sf1c7vp.pdf

In which Chapter is the Published Work: 12

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate:

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  Matt Williams designed the study, collected data, selected and conducted data analyses, and wrote the article.

Matt Williams Digitally signed by Matt Williams
DN: cn=Matt Williams, o=Massey University, ou=School of Psychology, email=m.n.williams@massey.ac.nz, c=NZ
Date: 2015.06.24 10:25:25 +12'00'

Candidate's Signature

24/6/2015

Date

Stephen Hill Digitally signed by Stephen Hill
DN: cn=Stephen Hill, o=Massey University, ou=School of Psychology, email=s.r.hill@massey.ac.nz, c=NZ
Date: 2015.06.24 15:45:32 +12'00'

Principal Supervisor's signature

24/6/2015

Date