
Prejudice toward Muslims in New Zealand:

Insights from the New Zealand Attitudes and Values Study

Chris G. Sibley^a, M. Usman Afzali^b, Nicole Satherley^a, Anastasia Ejoval^a, Samantha Stronge^a, Kumar Yogeeswaran^b, Michael Grimshaw^b, Diala Hawi^c, Zahra Mirnajafi^d, Fiona Kate Barlow^d, Petar Milojev^a, Lara M. Greaves^a, Sarah Kapeli^a, Elena Zubielevitch^a, Logan Hamley^a, Maria C. Basabas^a, Marvin H. Wu^a, Chloe Howard^a, Carol H. J. Lee^a, Yanshu Huang^d, Christopher Lockhart^a, Joaquín Bahamondes^a, Sam Manuela^a, Taciano L. Milfont^e, Ryan Perry^f, Nikhil K. Sengupta^g, Nickola C. Overall^a, John H. Shaver^h, Geoffrey Troughton^e, Danny Osborne^a, and Joseph Bulbulia^a

^a University of Auckland, NZ; ^b University of Canterbury, NZ; ^c Doha Institute for Graduate Studies, Qatar; ^d University of Queensland, Australia; ^e Waikato University, NZ; ^f University of Melbourne, Australia; ^g University of Kent, UK; ^h University of Otago, NZ

Following the March 15th Christchurch terrorist attack, members of our research team have been repeatedly asked to comment or provide summary statistics from the New Zealand Attitudes and Values Study (NZAVS) on prejudice toward Muslims. As the curators of the NZAVS, we think that these findings should be in the public domain and accessible to as wide an audience as possible. In this article, we aim to provide a comprehensive summary of what we know from the NZAVS about attitudes toward Muslims and prejudice in New Zealand more generally. From 2012 onwards, the NZAVS included a feeling thermometer rating of people's level of warmth toward Muslims. Here, we summarize what we know from the NZAVS about levels of warmth toward Muslims in the New Zealand population. We describe the distribution of thermometer ratings of warmth toward Muslims annually from 2012 onward, and compare these with thermometer ratings of a range of other groups that we also track. We present a regression model documenting the extent to which a broad range of demographics and aspects of personality are associated with low levels of warmth toward Muslims, and present a parallel model assessing warmth ratings toward immigrants as a comparison. Finally, we present a series of growth curve models outlining the relative level and rate of change over time in warmth toward Muslims and other groups from 2012-2018. Results from these analyses indicate that over the 2012-2018 period, levels of warmth toward Muslims in New Zealand were comparatively low relative to warmth ratings of other groups. However, warmth toward Muslims has also been steadily but gradually increasing over time in New Zealand.

Keywords: Prejudice, Muslim, Christchurch Terrorist Attack, Feeling Thermometer

Introduction

The shock and horror of the March 15th 2019 terrorist attack in Christchurch will shape our nation for many years to come. With 51 Muslim men, women, and children killed in the attack, 49 more injured, and many more deeply traumatised, this attack will indeed be remembered as one of the nation's darkest days. Our research team offer their heartfelt sympathies to the victims of the attack, their families, friends, colleagues and the broader Muslim community in Aotearoa, New Zealand. One of the authors (MUA), is a regular to the Al-Noor Mosque, and is among those grieving the loss of 14 friends, and also knows several others who were injured in the attacks.

In the days following the terrorist attack, members of our research team were repeatedly approached and asked

to comment or provide summary statistics on attitudes toward Muslims in New Zealand, levels of prejudice in New Zealand more generally, and to discuss how we think the attack has and will continue to shape our national identity. Central to such questioning is the placing of Muslims as not just a religious other but also an ethnic other; this exposes underlying attitudes and dualisms in our conceptions of national identity.

The reason our research team have been approached is because we collectively manage the New Zealand Attitudes and Values Study, or NZAVS. The NZAVS is a large-scale national probability longitudinal panel questionnaire study that surveys thousands of New Zealand citizens annually. Each year, our questionnaire asks New Zealanders to self-report on their attitudes toward a range of topics, their values, their level of

wellbeing, different aspects of their health, their personality, how they spend their time, and so forth. For those interested, more information about the NZAVS is available at: www.nzavs.auckland.ac.nz

The NZAVS questionnaire also includes a scale that asks people to rate their warmth toward different social groups. As part of this scale, since 2012, the NZAVS has measured self-reported ratings of warmth toward Muslims. In writing this paper, our goal is to present what we know from the NZAVS about attitudes toward Muslims in New Zealand in one place, in an accessible manner for as wide an audience as possible.

What is a feeling thermometer?

The measure of warmth toward Muslims that we include in the NZAVS questionnaire is generally known as a ‘feeling thermometer’, or just ‘thermometer scale.’ The thermometer scale was popularised in the American National Election Study, beginning in 1964. The scale was designed to provide a very simple and broad self-report instrument for measuring general positive (warm) versus negative (cold) feelings toward a range of different groups, topics, social policies, and presidential candidates (see Kinder & Drake, 2009, for discussion of thermometer scales used to track racism over time in the American

measures, including levels of agreement with attitude statements positioning historical injustices experienced by Māori as being no longer relevant, and attitude statements denying that Māori continue to experience discrimination.

Satherley and Sibley’s (2018) research provides a good example of how feeling thermometer ratings provide a window into the same overall generalized positive-negative attitude factor as other types of attitude statements assessing group attitudes in New Zealand. This is important because while feeling thermometers undeniably miss a lot of subtleties in people’s attitudes, they do reliably intercorrelate and load on the same overall attitude factor, or latent construct, as other more detailed and specific attitude statements (see also Bergh, Akrami, Sidanius, & Sibley, 2016, for additional research in this area). That is, people who rate a group less warmly on a feeling thermometer also tend to respond in similar, prejudicial ways to more specific questions about that group. This research thus provides evidence for the construct validity of feeling thermometers as measures of positive-negative ethnic group attitudes, or put simply, prejudice.

Research by Barlow and colleagues (2019) also provides good evidence for the construct validity of the

feeling thermometers employed in the NZAVS (see also Perry, Priest, Paradies, Barlow & Sibley, 2018, for additional evidence of the validity of these scales in the New Zealand context). Barlow et al. demonstrated that how people scored on feeling thermometers measuring their

Please rate your feelings of WARMTH toward the following groups using the “feeling thermometer scale” for each group.

Feel LEAST WARM Toward This Group	Neutral						Feel MOST WARM Toward This Group								
1	2	3	4	5	6	7									
NZ Europeans	1	2	3	4	5	6	7	Overweight people	1	2	3	4	5	6	7
Māori	1	2	3	4	5	6	7	Immigrants in general	1	2	3	4	5	6	7
Asians in general	1	2	3	4	5	6	7	Chinese	1	2	3	4	5	6	7
Pacific Islanders	1	2	3	4	5	6	7	Indians	1	2	3	4	5	6	7
Elderly people	1	2	3	4	5	6	7	Muslims	1	2	3	4	5	6	7
Refugees	1	2	3	4	5	6	7	People with mental illness	1	2	3	4	5	6	7

Figure 1. The feeling thermometer scale included in the Time 9 (2017) wave of the NZAVS questionnaire.

National Election Study).

An exact copy of the thermometer scale used in the most recent (2017) wave of the NZAVS questionnaire is presented in Figure 1. As you can see, the NZAVS version of the scale simply asks people to rate their warmth toward a range of different groups from 1 (feel least warm) to 7 (feel most warm). The specific scale wording is based on earlier psychometric work that we conducted assessing warmth ratings toward a much broader range of groups in New Zealand. The evidence from this early work indicates that the NZAVS thermometer scale provides a valid and reliable index of overall generalized positive-negative attitudes toward a large range of different groups (see Duckitt & Sibley, 2007, for information about scale validation in New Zealand).

Feeling thermometer research in New Zealand

Feeling thermometers have been used to test a fairly broad range of research questions relating specifically to prejudice and intergroup attitudes in the New Zealand context. Most of this work is from our research group and uses data from the NZAVS. Satherley and Sibley (2018), for example, validated a new measure of modern racism toward Māori. They showed that thermometer ratings assessing negative affect toward Māori loaded on the same superordinate or overall factor as a variety of other

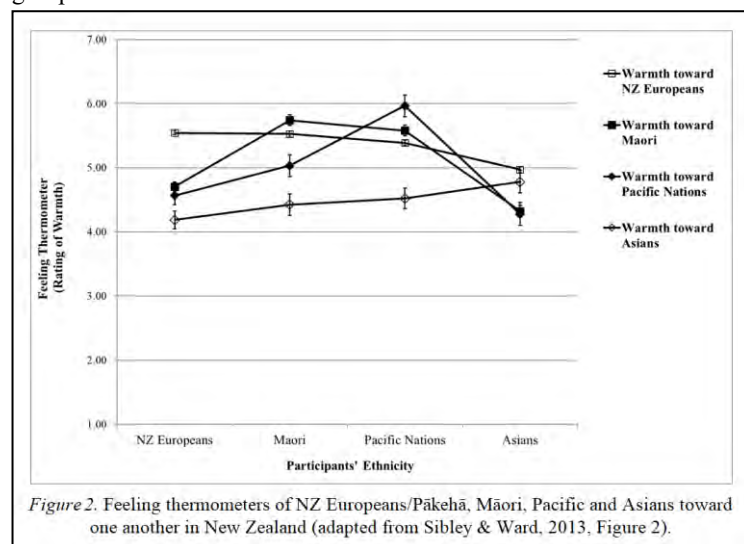
warmth and anger toward European/Pākehā, Māori, Pacific and Asian people was predicted by their level of contact with those same groups. The contact hypothesis is somewhat of a cornerstone theory of prejudice reduction in social psychology (see, for example, Pettigrew & Tropp, 2006). The theory states that one way to reduce prejudice toward different groups is to have members of the groups come into contact with one another under positive conditions, ideally with shared goals, and in ways supported by the broader community. Barlow and colleagues’ (2019) findings replicate and extend an extremely large body of previous research from around the world studying how intergroup contact reduces prejudice and shows that the widely observed effect of intergroup contact holds when measuring prejudice using short and simple feeling thermometers (see also Sengupta, Barlow, & Sibley, 2012, for additional research on contact and feeling thermometers using NZAVS data).

Barlow and colleagues (2019) also showed that feeling thermometers of warmth and anger were predicted by different types of contact with other ethnic groups: positive contact with members of another ethnic group predicted increased feelings of warmth toward them over time but predicted a much weaker reduction in anger.

Negative contact, in contrast, led to increased feelings of anger toward them, but predicted a much weaker reduction in warmth. This is also an important finding in its own right because it shows that merely increasing how much members of groups interact with one another might have unintended negative effects if some of the contact that occurs is negative or when groups are in competition with one another.

We emphasise again that although feeling thermometers certainly miss nuance in people's attitudes, they remain extremely useful when one wants to conduct comparative analyses assessing how positive or negative people feel toward different social groups. The measurement of simple warm-cold thermometer ratings toward a range of groups allows researchers to ask questions such as 'which group is most liked or disliked?' and 'how warm do people feel toward their own group relative to other groups?'

Sibley and Ward (2013), for example, analysed feeling thermometer ratings from the first (2009) wave of the NZAVS. They conducted what one might call a '360 degree' analysis of NZ European/Pākehā, Māori, Pacific and Asian participant's ratings of warmth toward their own group, and also their levels of warmth toward each of the three other groups. Unsurprisingly, Sibley and Ward showed that each ethnic group tended to express fairly high levels of warmth toward their ingroup (or ingroup bias). However, the results also revealed a series of fairly subtle asymmetries in how each group felt toward the other groups—it was certainly not the case that people merely liked their own group more and expressed the same magnitude of reduced warmth toward other ethnic groups.



The '360 degree' analysis reported by Sibley and Ward (2013) is reproduced in Figure 2. As can be seen in Figure 2, NZ European/Pākehā, Māori, and Pacific people all expressed the lowest level of warmth toward Asian peoples. Māori, Pacific and Asian people, by contrast, all expressed relatively high warmth toward NZ

European/Pākehā. Pacific people also expressed a high level of warmth toward Māori, which Sibley and Ward suggested was consistent with Teaiwa and Mallon's (2005) notion of ambivalent kinship between the two groups.

Sibley and Ward's (2013) analysis only included European/Pākehā, Māori, Pacific and Asian peoples due to the small number of participants from other ethnic groups in the sample (e.g., Middle Eastern, Latin American or African). Thus, although it is easily possible to look at warmth toward all other groups included in the thermometer scale (we report this data in full in the current paper), it was not possible to conduct a full '360 degree' analysis including responses from a larger number of ethnic groups due to limited sample size.

Hamley, Houkamau, Osborne, Barlow, and Sibley (2020) built upon Sibley and Ward (2013) by presenting a more focused examination of Māori and NZ European/Pākehā ratings of warmth toward both their own group (or ingroup) and each other's group (or outgroup). Using Latent Profile Analysis, Hamley and colleagues (2020) showed that some Māori and NZ European/Pākehā expressed a specific combination of high ingroup, but low outgroup warmth, which is indicative of ingroup bias. However, this group was made up of a relatively small number of people in both cases (about 7% of Māori and 10% of NZ Europeans). The majority of both Māori and NZ European/Pākehā scored the feeling thermometers in a consistent pattern indicating high warmth toward their ingroup, but also high warmth toward each other's group (about 59% of Māori and 54% of NZ Europeans).¹

Most of the research on feeling thermometers

(including the results we report in the current paper) examine overall mean or proportional differences in feeling thermometers. The research by Hamley et al. (2020) is particularly important because it clearly shows that just examining means and the distribution around them is not necessarily the full story. If research only examines overall mean (or median, or mode, or summed) ratings, then there is a risk of missing subtler differences in the distinct combinations of high/moderate/low warmth that people hold across multiple groups. These more subtle patterns are not contradictory to research examining mean differences (statistically speaking), but a lot can be missed when only examining overall aggregate summaries, or assuming that there is only one underlying distribution to people's scores in the population, rather than many distinct types of people who express qualitatively distinct high/moderate/low levels of warmth or positive emotion toward different groups (see also Osborne & Sibley, 2017, for a general discussion of this topic in political science).

¹ Hamley and colleagues (2020) also showed that even the psychological factors underlying why a small minority fit an ingroup bias profile differed for Māori and NZ European/Pākehā. For Māori, ingroup bias seems to be strongly driven by the motivation to maintain strong and coherent ingroup identity; for NZ

European/Pākehā, it seems to be more strongly driven by Social Dominance Orientation (Sidanius & Pratto, 1999). This is an orientation based on the competitive-driven goal of group-based dominance and superiority.

Warmth toward Muslims in New Zealand: What do we already know?

As far as we are aware, prior to the March 15th terrorist attack, there had been only two published quantitative studies examining feeling thermometer ratings of Muslims (or other quantitative scales assessing anti-Muslim or anti-Islamic sentiment) in New Zealand. Both studies used data from the NZAVS and both examined feeling thermometer ratings (which at this time is the only measure of attitudes toward Muslims in the NZAVS).² Readers are directed to see also Highland, Troughton, Shaver, Barrett, Sibley and Bulbulia (2019); Hawi, Osborne, Bulbulia and Sibley (2019); and Yogeewaran, Afzali, Andrews, Chivers, Wang, Devos and Sibley (2019) that also analyzed ratings of warmth toward Muslims using NZAVS data.

Shaver, Troughton, Sibley and Bulbulia (2016) used data from the 2013 (Time 5) wave of the NZAVS to analyse feeling thermometer ratings of Muslims, Arabs, and ‘immigrants in general’. Shaver and colleagues noted a fairly large difference in levels of warmth toward these groups, with warmth ratings of both Muslims and Arabs being near-identical and markedly lower (.5 units on the scale from 1-7) than warmth toward immigrants. The goal of Shaver and colleagues’ research was to test specific predictions from Evolutionary Lag Theory. This theory broadly states that under specific conditions (such as in New Zealand) religion should predict increased levels of tolerance toward other religious groups. Shaver et al. (2016) found good support for this prediction by showing that among Christians, a higher level of religious identification was associated with higher levels of warmth toward Muslims. In the wake of the March 15th terrorist attack, this finding is particularly relevant, and shows that the Christian community in New Zealand may help to support norms of acceptance, inclusion and warmth toward Muslims. Interfaith gatherings and support offered to the Muslim community after the attack also illustrate this point.

A second study by Shaver, Sibley, Osborne and Bulbulia (2017) also examined feeling thermometer ratings of Muslims using data from the 2013 (Time 5) wave of the NZAVS. Shaver et al. examined the association between rates of overall media exposure (self-reported hours of news consumption per week) and feeling thermometer ratings of warmth and anger toward Muslims (see also Kabir & Bourk, 2012, for qualitative analysis of media representation of Islam and Muslims in New Zealand media). Results indicated that hours of news consumption was associated with lower warmth and increased anger toward Muslims specifically, but not warmth and anger toward Asians (the ethnic group to

which most of New Zealand’s and the world’s Muslims belong). Moreover, these associations held when adjusting for a broad range of other factors, such as education, age, gender, socio-economic status, and political orientation.

The lack of moderating effect for political orientation observed by Shaver et al. (2017) is particularly noteworthy, as it did not interact with hours of news exposure. This suggests that the news New Zealanders consume has a general effect of decreasing warmth toward Muslims, rather than being driven by those in a specific liberal or conservative news bubble. This, in turn, suggests that a general negative representation of Muslims and Islam may span much of the news, rather than reflecting a specific partisan news bias. The lack of similar effects observed for warmth toward Asians further suggests that this news bias is specific and localized anti-Muslim sentiment.

Outside of the NZAVS, there has been very little recent research on prejudice towards Muslims in New Zealand. An exception is Greaves et al. (2020), who conducted a nationally representative survey on religion as part of the International Social Survey Programme in the months before the March 15th attack. Participants rated Muslims more negatively, and as far more threatening, than they did other religious groups such as Christians, Hindus, Buddhists, Jews, and ‘atheists or non-believers’. Negativity toward Muslims and other religious groups was measured in both 2008 and in 2018, and had decreased over the time period reflecting a small reduction in prejudice. The percentage of people who agreed that Muslims were ‘somewhat’ or ‘very’ threatening had decreased from 22.5% in 2008 to 19.5% in 2018. The predictors of prejudice were similar to those found in past studies using the NZAVS (e.g. Shaver et al., 2016, 2017). These findings also support work from Wilson (2019) who found that participants from a convenience sample found Muslims more threatening than other religious groups. This past research shows that the results of the NZAVS likely generalise when also comparing religious groups with other religious groups, rather than with ethnicities (as is the case with the NZAVS).

Overview of the present study

The two papers by Shaver and colleagues (2016, 2017) attracted widespread interest following the March 15th terrorist attack. However, the goal of these two papers was to test specific, fine-grained theoretical predictions, rather than comprehensively and broadly describe levels of warmth toward Muslims in New Zealand. In the current research, we first aim to address this need from the community by comprehensively describing the

² There is also one other paper by Duckitt and Sibley (2007) which measured warmth toward the social category labelled ‘Arabs’, in a sample of 212 undergraduate students. Duckitt and Sibley (2007) measured warmth toward a larger number of different groups in their study, and showed that variation in peoples levels of warmth toward Arabs was weakly correlated with Right-Wing Authoritarianism (an individual difference measure that indexes the threat-driven motivation for collective security and social cohesion) and more strongly correlated with Social Dominance Orientation (an individual difference measure that indexes the competitive-driven goal of group-based dominance and superiority). The paper by Duckitt and Sibley (2007) shows that, together, Right-Wing

Authoritarianism and Social Dominance Orientation explain many different forms of prejudice, and prejudice toward many different groups (see also Duckitt, 2001, for an excellent discussion of this topic). Their research suggests that back in 2002 when these data were collected, and at least among undergraduate university students in New Zealand, negative feelings toward Arabs were more strongly linked with the desire for group hierarchy and dominance over groups seen as weak and inferior (i.e., Social Dominance Orientation) than they were by the desire for security and cohesion based on perceived danger and threat (i.e., Right-Wing Authoritarianism).

distribution, means and frequencies for feeling thermometer data assessing warmth (and anger) toward Muslims from all waves of the NZAVS which have included these measures. We also include similar descriptive data for ratings of ethnic groups for which we collected feeling thermometer ratings as a point of comparison.³

Second, we present a regression model assessing the extent to which a broad range of demographics and aspects of personality are associated with concurrent levels of warmth toward Muslims in the latest (2017) wave of the NZAVS. Similar to the comparison analysis of Asians included in the news consumption analysis by Shaver et al. (2017), we present a parallel model assessing the extent to which the same set of demographics and personality traits are associated with warmth toward immigrants. This model is useful because it comprehensively profiles the various demographics associated with feelings of lesser or greater warmth toward Muslims in New Zealand using the most recent available NZAVS data.

Third, we directly answer the question of whether levels of warmth toward Muslims are increasing or decreasing over time in New Zealand. We examine change in levels of warmth toward Muslims from 2012-2018, as the NZAVS only began including a thermometer rating of Muslims in 2012. We model rates of change in warmth using Latent Growth Curve analysis. Because the NZAVS is a longitudinal panel survey, the study tracks the same people consistently each year. Our analysis of change therefore describes the average rate of change within individuals over time. This is an important, but subtle, point, and should not be confused with analyses of means or percentages from repeated annual cross-sectional samples (such as the International Social Survey Programme data reported in Greaves et al., 2020), which randomly sample different groups of people each year. Analyses of these latter types of data can provide useful data on trends in the overall population level of something over time. It is crucial to realize that they do not actually test change within individuals over time, however. This is an important limitation because it is often change within individuals that is actually of interest.⁴ Again, we also report growth curves describing how feeling thermometer ratings of other immigrant groups ('Chinese', 'Indians', 'Asians in general' and 'Immigrants in general') have changed over this same time period as a comparison.⁵

METHODS

Participants

Demographics for the Time 4 (2012) – Time 9 (2017)

³ The NZAVS questionnaire has not, up to this point, included thermometer ratings of other religious groups.

⁴ Consider the following simple example. Imagine you collect two waves of data a year apart, from the same 5 people each year. Each person's scores at year 1 and again at year 2 are as follow: person one: 1,2; person two: 4,5; person three: 9,10; person four: 8,9; person five: 9,2. If you take the average of these 5 scores then you observe a mean of 6.2 at year 1 and a mean of 5.6 at year 2. You conclude that people's scores are going down by about .6 units per year. This is incorrect. The population average may have decreased, but 4 of the 5 people are actually increasing by one unit each year. There is just one person that has shown a decrease, albeit a

dramatic one. Of course, the bias can be subtler than this. This example merely shows how one can draw the wrong conclusions depending on the level of analysis in a very simple case.

Sampling Procedure

The original Time 1 (2009) NZAVS sample was drawn from a random sample of the New Zealand Electoral Roll. Postal questionnaires were sent to 40,500 registered voters or roughly 1.36% of all registered voters in New Zealand. The overall response rate (adjusting for the address accuracy of the Electoral Roll and including anonymous responses) was 16.6%. To boost sample size at Time 3 (2011) and compensate for sample attrition, a booster sample was recruited through an unrelated survey posted on the website of a major New Zealand newspaper in 2011. A total of 3,208 participants registered an initial expression of interest in being contacted to participate in the NZAVS via this survey. Participants in this non-random booster sample were emailed an invitation to participate in an online version of the NZAVS, and those who did not respond to the email were also sent a postal version of the questionnaire. A total of 2,962 participants completed the questionnaire when subsequently contacted (92.4%).

Full details about the sampling procedure, response rates and so on for the 2009-2011 waves of the NZAVS are provided in Sibley (2019a). Full sampling details for the waves analysed in the current paper, from 2012-2017 (Times 4-9), are included in the Appendix. Thermometer ratings of Muslims were included in the NZAVS from Time 4 (2012) onwards (hence, we do not report on data from before this time point in this paper).

Measures

Feeling thermometer ratings were measured using the following question: "Please rate your feelings of warmth toward the following groups using the 'feeling thermometer scale' for each group". The rating scale ranged from 1 (Feel LEAST WARM toward this group) through 4 (Neutral) to 7 (Feel MOST WARM toward this group). An exact copy of the question wording and formatting from the Time 9 NZAVS questionnaire is presented in Figure 1.

⁵ These are the labels that are used in the feeling thermometer rating scales in the NZAVS, hence their use in quotes. We also note that Muslims as a group include a wide diversity of cultures and ethnicities and it is indicative of contemporary global discourse that one religious identity very often becomes collapsed into a type of popular and populist ethnicized and collectivist immigrant discourse. The NZAVS research group are aware of this issue and this is indeed central to our research regarding warmth towards Muslims.

Table 1. Sample details for the Time 4 – Time 9 waves of the NZAVS.

Time (Year):		4 (2012)	5 (2013)	6 (2014)	7 (2015)	8 (2016)	9 (2017)
N		12,179	18,261	15,820	13,942	21,936	17,072
Gender	Women	7,621	11,460	10,002	8,726	13,722	10,802
	Men	4,554	6,798	5,798	5,197	8,171	6,246
	Diverse	-----	-----	20	17	41	24
Ethnicity	Europeans	10,281	15,607	14,161	12,506	19,478	15,606
	Māori	2,025	2,328	1,977	1,672	2,473	2,007
	Pacific Nations	610	625	526	427	574	466
	Asian	619	814	684	545	1,001	700
Age	Mean(SD)	49.11(15.02)	47.65(14.07)	49.33(14.03)	50.79(13.90)	49.62(13.93)	51.33(13.77)
Education¹	Mean(SD)	4.70(2.84)	4.93(2.82)	5.08(2.81)	5.19(2.78)	5.30(2.75)	5.26(2.77)
Income²	Mean(SD)	11.23(0.80)	11.30(0.79)	11.25(1.14)	11.30(0.97)	11.29(1.04)	11.37(0.85)
Deprivation³	Mean(SD)	5.02(2.85)	4.81(2.79)	4.71(2.76)	4.70(2.78)	4.67(2.75)	4.59(2.72)
SES⁴	Mean(SD)	52.30(16.27)	52.94(16.46)	53.73(16.26)	54.36(16.04)	54.30(16.29)	54.88(16.16)
Categorical Demographics	Religious	4,934	6,879	6,030	5,844	8,274	6,137
	Parent	8,856	13,084	11,655	10,511	15,883	12,628
	Partner	8,314	12,984	11,354	10,090	15,885	12,770
	Employed	8,464	13,578	12,037	10,457	17,119	13,136
	Urban	7,871	12,151	10,454	8,994	14,074	13,732
	NZ Born	9,421	14,225	12,497	10,624	17,157	13,517

¹ NZReg 0-10; ² Household income (Log); ³ 2013 Meshblock NZDep 1-10; ⁴ 2013 Socioeconomic Index SEI 10-90

From Times 4-6 the NZAVS questionnaire also included thermometer ratings assessing anger toward these same groups. The anger thermometer used the following instructions: ‘Now please rate your feelings of anger toward the same groups using the scale below.’ The rating scale ranged from 1 (Feel NO ANGER toward this group) through 4 (Neutral) to 7 (Feel ANGER toward this group). The anger thermometer followed immediately after the warmth thermometer in the questionnaire.

Big-Six personality was measured using the Mini-IPIP6 (Sibley et al., 2011). Each trait is measured using 4 marker items, rated from 1 (very inaccurate) to 7 (very accurate) and averaged to give scale scores for Extraversion ($\alpha = .76$), Agreeableness ($\alpha = .72$), Conscientiousness ($\alpha = .69$), Neuroticism ($\alpha = .74$), Openness to Experience ($\alpha = .71$), and Honesty/Humility ($\alpha = .77$). Internal reliability estimates are reported for the Time 9 wave, but are similar in all waves.⁶

With regard to demographics, deprivation was assessed using the 2013 New Zealand Deprivation Index, which uses census information to assign a decile-rank index from 1 (least deprived) to 10 (most deprived) to each meshblock unit (Atkinson, Salmond, & Crampton, 2014). Participants’ socio-economic index was measured using the occupation-based rating developed by Milne, Byun, and Lee (2013). This index assigns a score from 10 to 90, where 90 indicates high socio-economic status. Education was coded into an eleven-level ordinal variable following the NZQA education coding scheme (0 = no qualification to 10 = doctorate).

The complete data dictionary listing all questions included in each wave of the NZAVS is available in

Sibley (2019b). A full copy of all NZAVS questionnaires is provided in Sibley (2019c). Details on post-stratification sample weighting are provided in Sibley (2019d).

RESULTS

Descriptive statistics

Tables 2-4 present means, standard deviations, and distributional information (percentages of extreme ratings) for thermometer ratings of warmth and anger toward different groups from Time 4 (2012) to Time 9 (2017). These descriptive statistics applied a post-stratification sample weighting to correct for sampling bias in gender, ethnicity and region in each wave of the study (see Sibley, 2019b).⁷

Histograms displaying the distribution of thermometer ratings of warmth and anger toward each group for each wave of the NZAVS from Time 4 (2012) to Time 9 (2017) are presented in Figures 3-8. As with Tables 2-4, these figures provide warmth ratings for all years, and anger ratings for the years where they were also measured. The numbers on top of each bar represent the percentage of people who rated that point on the thermometer scale. The data presented in these histograms were also weighted following Sibley (2019b), and hence represent estimates of the distribution of warmth and anger in the population of New Zealand adults. Note that the percentages in these tables and figures represent those of valid responses (i.e., missing data are not included in the denominator). For each feeling thermometer rating we also re-estimated the histogram excluding members of that group from the analysis in order to estimate how others in New Zealand felt toward each group.

⁶ The Big-Six personality model provides, in our view, the most representative and valid representation of the broad structure of individual differences in personality. The theory extends earlier Big-Five models of personality by identifying an additional dimension of Honesty-Humility, and is supported by a vast range of studies, across numerous cultures (although it is more robust in Western nations; see Ashton & Lee, 2007). The Mini-IPIP6 Big-Six measure of Big-Six personality has been extensively validated in New Zealand, its factor structure is reliable and it shows good convergent and discrimination validity (Sibley et al., 2011), is stable over time (Milojev, Osborne, Greaves, Barlow & Sibley,

2013), has acceptable item response properties (Sibley, 2012), and demographic norms are available (see Sibley & Pirie, 2013).

⁷ Excluding people who identify as Muslim from the data did not change this overall pattern, and the same was the case when people reporting Middle Eastern ethnicity were removed from frequency calculations for warmth and anger towards Arab people (in the tables and graphs, these patterns are referred to as patterns of ratings by the “outgroup”).

Table 2. Descriptive statistics for thermometer ratings of 'Muslims', 'Arabs' and 'Indians' from the NZAVS.

Question	Year	Weighted mean	95%CI low	95%CI high	Weighted SD	% of people resp. 1	% of people resp. 7	Weighted N
'Muslims'								
Warmth	2012	3.65	3.61	3.68	1.55	12.16	4.21	11,785
	2013	3.74	3.71	3.76	1.57	10.97	5.04	17,462
	2014	3.84	3.81	3.87	1.55	9.87	4.74	15,075
	2015	3.80	3.77	3.83	1.54	10.45	3.94	13,548
	2016	3.91	3.88	3.93	1.54	9.22	4.84	21,095
	2017	4.04	4.01	4.06	1.48	7.40	4.79	16,563
Anger	2012	3.01	2.98	3.05	1.85	33.43	5.44	11,778
	2013	2.97	2.93	3.00	1.83	33.59	5.20	17,424
	2014	3.09	3.05	3.12	1.82	29.90	5.19	15,157
'Arabs'								
Warmth	2012	3.73	3.71	3.76	1.49	9.04	4.30	11,780
	2013	3.78	3.76	3.81	1.50	8.45	4.84	17,428
	2014	3.88	3.85	3.91	1.49	7.73	4.62	15,079
	2015	3.82	3.79	3.85	1.46	8.07	3.83	13,515
	2016	3.88	3.86	3.90	1.48	7.52	4.53	21,017
	2017	-	-	-	-	-	-	-
Anger	2012	2.93	2.89	2.96	1.79	34.35	3.97	11,752
	2013	2.89	2.86	2.91	1.75	34.27	3.69	17,421
	2014	2.97	2.93	3.00	1.73	31.47	3.31	15,161
'Indians'								
Warmth	2012	4.12	4.09	4.15	1.41	4.93	5.62	11,796
	2013	4.21	4.18	4.23	1.42	4.47	6.49	17,488
	2014	4.32	4.30	4.35	1.40	3.55	6.37	15,119
	2015	4.23	4.20	4.25	1.37	3.89	4.68	13,558
	2016	4.23	4.21	4.25	1.42	4.45	5.67	21,135
	2017	4.25	4.23	4.28	1.39	3.99	5.42	16,568
Anger	2012	2.67	2.64	2.70	1.62	37.46	2.05	11,782
	2013	2.66	2.63	2.69	1.64	37.90	2.29	17,463
	2014	2.73	2.70	2.76	1.60	34.71	1.83	15,192

We present these data for broadly descriptive purposes and so that the full set of information from the NZAVS on feeling thermometer ratings is available for other researchers, government departments, and other community groups. More detailed descriptive information, including full information of the distribution of scores used to create the histograms in Figures 3-8, is available in an online supplement at www.nzavs.auckland.ac.nz on the NZAVS bibliography page.

Demographic Differences in Warmth toward Muslims

We conducted a multiple regression predicting warmth toward Muslims using a range of demographic and personality factors ($N = 16,641$). Warmth towards 'immigrants in general' was also included as a dependent

models are presented in Table 6. Missing data for exogenous variables were estimated using Rubin's (1987) procedure for multiple imputation with parameter estimates averaged over 2,000 datasets (thinned using every 100th iteration). The model was estimated with Maximum Likelihood with robust estimation of standard errors.

As can be seen in Table 6, lower warmth toward Muslims was predicted by being male, being older, lower educational attainment, lower socioeconomic status, being religious, not being Muslim, being unemployed, and living in rural areas. There was no significant association between warmth toward Muslims and household income, deprivation levels, ethnicity, having children or a partner, or being born in New Zealand. In terms of personality variables, lower warmth toward Muslims was predicted by lower Extraversion,

variable so as to provide a comparison group. Bivariate correlations between all variables included in the model are reported in Table 5. The results from these regression

Agreeableness, Openness, and Honesty-Humility, and higher Conscientiousness and Neuroticism.

Table 3. Descriptive statistics for thermometer ratings of 'Chinese', 'Asians in general' and 'Pacific Islanders' from the NZAVS.

Question	Year	Weighted mean	95%CI low	95%CI high	Weighted SD	% of people resp. 1	% of people resp. 7	Weighted N
'Chinese'								
Warmth	2012	4.27	4.25	4.30	1.36	3.36	6.08	11,803
	2013	4.35	4.33	4.38	1.37	3.04	7.12	17,486
	2014	4.44	4.42	4.47	1.37	2.85	7.08	15,113
	2015	4.34	4.32	4.37	1.35	3.22	5.59	13,558
	2016	4.33	4.30	4.35	1.39	3.59	6.04	21,119
	2017	4.36	4.33	4.38	1.35	2.81	5.85	16,574
Anger	2012	2.62	2.59	2.65	1.58	37.71	1.65	11,783
	2013	2.59	2.56	2.62	1.58	38.59	1.63	17,448
	2014	2.70	2.67	2.73	1.58	34.95	1.51	15,192
'Asians in general'								
Warmth	2012	4.47	4.44	4.49	1.33	2.55	7.46	11,802
	2013	4.56	4.53	4.58	1.33	2.12	8.19	17,469
	2014	4.62	4.59	4.64	1.32	2.10	7.63	15,127
	2015	4.53	4.50	4.55	1.28	2.07	5.98	13,531
	2016	4.54	4.52	4.56	1.32	2.07	6.94	21,086
	2017	4.55	4.52	4.57	1.29	1.83	6.72	16,590
Anger	2012	2.61	2.58	2.64	1.58	38.14	1.67	11,787
	2013	2.57	2.54	2.60	1.59	39.41	1.74	17,468
	2014	2.68	2.65	2.71	1.58	35.33	1.44	15,219
'Pacific Islanders'								
Warmth	2012	4.67	4.65	4.70	1.32	1.49	9.86	11,800
	2013	4.74	4.72	4.77	1.34	1.70	10.92	17,492
	2014	4.82	4.79	4.84	1.30	1.30	10.55	15,132
	2015	4.74	4.71	4.76	1.27	1.51	8.22	13,533
	2016	4.79	4.76	4.81	1.31	1.50	10.03	21,111
	2017	4.77	4.75	4.80	1.28	1.48	9.27	16,586
Anger	2012	2.57	2.54	2.61	1.59	39.37	1.81	11,798
	2013	2.52	2.49	2.54	1.59	41.08	1.88	17,464
	2014	2.61	2.58	2.64	1.55	37.02	1.32	15,221

As with warmth toward Muslims, lower warmth toward immigrants in general was predicted by being male, lower education, lower socioeconomic status, and living in rural areas. For personality variables, the same pattern was identified for warmth toward immigrants as for warmth toward Muslims, suggesting that personality traits predict prejudice more generally (i.e., the associations are not specific to attitudes toward Muslims).

There were some differences in the pattern of predictors for warmth toward Muslims as compared to warmth toward 'immigrants in general'. Lower warmth toward immigrants was uniquely predicted by lower household income, Māori ethnicity, being a parent, and being born in New Zealand, while higher warmth toward immigrants was predicted by Pacific ethnicity. These variables were unassociated with warmth toward Muslims. In contrast, lower warmth toward Muslims was uniquely predicted by being older, not being Muslim, and being unemployed; these variables were not associated

with warmth toward immigrants.

Changes in Warmth toward Muslims from 2012-2018

We assessed changes in warmth toward Muslims in New Zealand from 2012-2018 (Time 4 – Time 9 of the NZAVS) using Latent Growth Modelling (LGM). LGM allows us to examine change in warmth over time *within* individuals, rather than change resulting from changes in sample composition over time. The LGMs presented here examine the same group of individuals over time, estimating a distinct latent trajectory of change over time for each individual in the analysis. These trajectories are then averaged to create an overall image of within-person change in the population. Although we focus on changes in warmth toward Muslims, we also estimated growth curves for warmth ratings of 'Asians in general', 'Immigrants in general', 'Chinese' and 'Indians', which provides a useful reference as to how attitudes toward

Table 4. Descriptive statistics for thermometer ratings of 'Immigrants in general', 'NZ Europeans' and 'Māori' from the NZAVS.

Question	Year	Weighted mean	95%CI low	95%CI high	Weighted SD	% of people resp. 1	% of people resp. 7	Weighted N
'Immigrants in general'								
Warmth	2012	4.43	4.41	4.46	1.25	1.96	6.32	11,802
	2013	4.49	4.47	4.51	1.26	1.90	7.04	17,468
	2014	4.55	4.53	4.57	1.24	1.61	6.66	15,098
	2015	4.47	4.44	4.49	1.23	1.85	5.38	13,548
	2016	4.51	4.49	4.53	1.27	1.88	6.51	21,090
	2017	4.51	4.49	4.54	1.25	1.93	6.33	16,563
Anger	2012	2.58	2.55	2.61	1.54	38.33	1.29	11,787
	2013	2.58	2.55	2.60	1.55	38.47	1.46	17,440
	2014	2.67	2.64	2.70	1.53	35.10	1.16	15,184
'NZ Europeans'								
Warmth	2012	5.53	5.50	5.55	1.25	0.57	27.99	11,812
	2013	5.58	5.55	5.60	1.28	0.68	30.74	17,512
	2014	5.60	5.58	5.63	1.23	0.31	29.70	15,127
	2015	5.52	5.49	5.54	1.23	0.55	26.18	13,543
	2016	5.56	5.54	5.58	1.24	0.47	28.37	21,114
	2017	5.51	5.49	5.54	1.24	0.38	27.27	16,570
Anger	2012	2.31	2.27	2.34	1.60	48.85	1.93	11,799
	2013	2.27	2.25	2.30	1.61	50.15	2.10	17,485
	2014	2.41	2.39	2.44	1.59	44.12	1.56	15,234
'Māori'								
Warmth	2012	4.85	4.82	4.88	1.37	1.55	13.14	11,810
	2013	4.92	4.90	4.95	1.37	1.69	14.51	17,498
	2014	5.00	4.98	5.03	1.32	1.34	13.67	15,136
	2015	4.93	4.91	4.96	1.30	1.31	11.88	13,538
	2016	5.00	4.98	5.02	1.31	1.20	13.45	21,091
	2017	4.99	4.96	5.01	1.29	1.24	12.76	16,593
Anger	2012	2.69	2.66	2.72	1.70	37.74	2.57	11,788
	2013	2.63	2.60	2.66	1.69	39.29	2.55	17,479
	2014	2.69	2.66	2.72	1.64	35.71	1.87	15,235

Muslims specifically may be changing relative to attitudes toward other groups in New Zealand who also tend to be viewed less favourably.

Participants provided ratings of warmth at each year (or wave) of data collection (Time 4 – Time 9). However, rather than fixing participants' responses to a single year, we allowed time scores (the time each individual's rating of warmth was made each year) to vary between individuals within each year to match the specific day of the year an individual's response was received. The model intercept for time (i.e., a score of 0) reflected the mean level of warmth expressed by participants in October 2012 (the start of data collection for Time 4). As such, we were able to estimate mean warmth at any point in time (in this

case, every three months), rather than once for each year.⁸

For completeness, we ran models that included linear, quadratic, and cubic terms for time. If necessary, these models were then re-run after removing higher order terms (e.g., cubic, then quadratic terms) when they were not significant. Analyses were limited to participants who completed at least 4 of the 6 waves, with missing data among participants who completed 4 or more waves estimated using Full Information Maximum Likelihood, with data assumed to be missing at random. Individual trajectories of change were weighted, such that individuals who completed more time points contributed the most information to the estimated mean trajectory of change.

⁸ Note that data collection for each wave of the NZAVS tends to run from October of one year to October of the next year, with data collection distributed throughout the 12 months. Thus, although the Time 9 NZAVS is referred to for simplicity as the 2017 wave, data

collection for that wave spanned October 2017 to October 2018. This is why we estimate model-implied rates of change for our LGMs from October 2012 – October 2018.

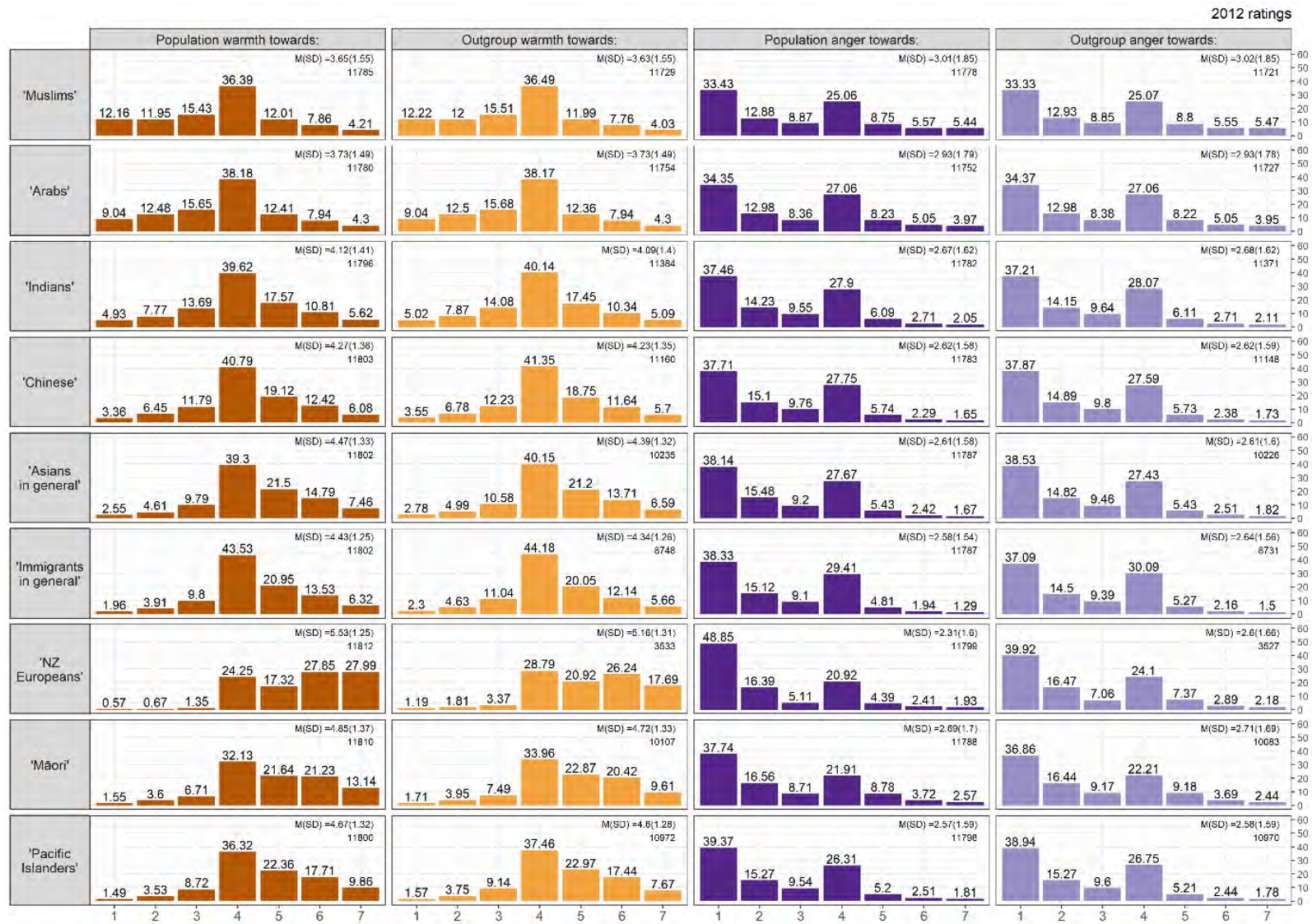


Figure 3. Histograms showing the distribution of thermometer ratings of warmth and anger toward each group in the Time 4 (2012) wave of the NZAVS.

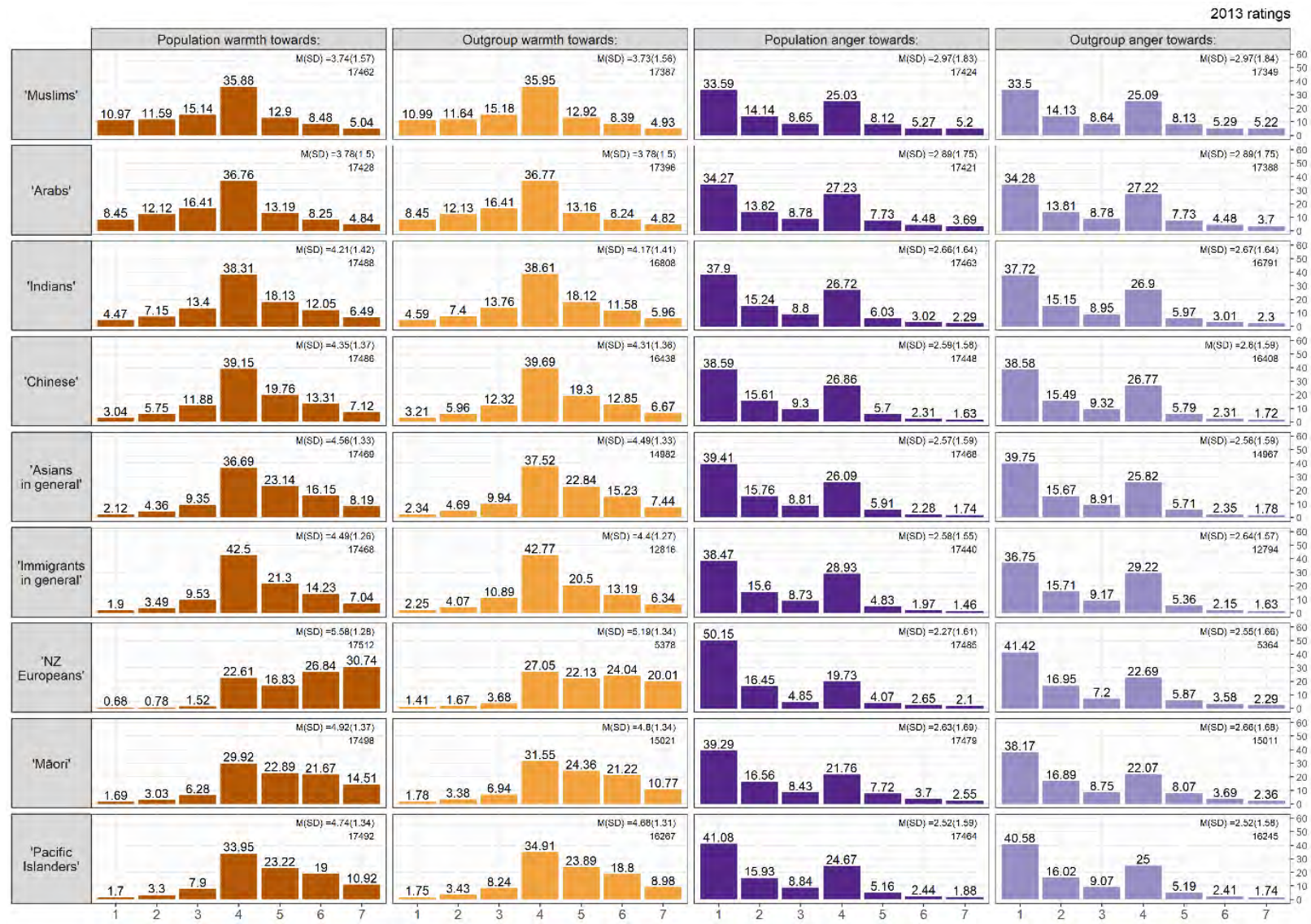


Figure 4. Histograms showing the distribution of thermometer ratings of warmth and anger toward each group in the Time 5 (2013) wave of the NZAVS

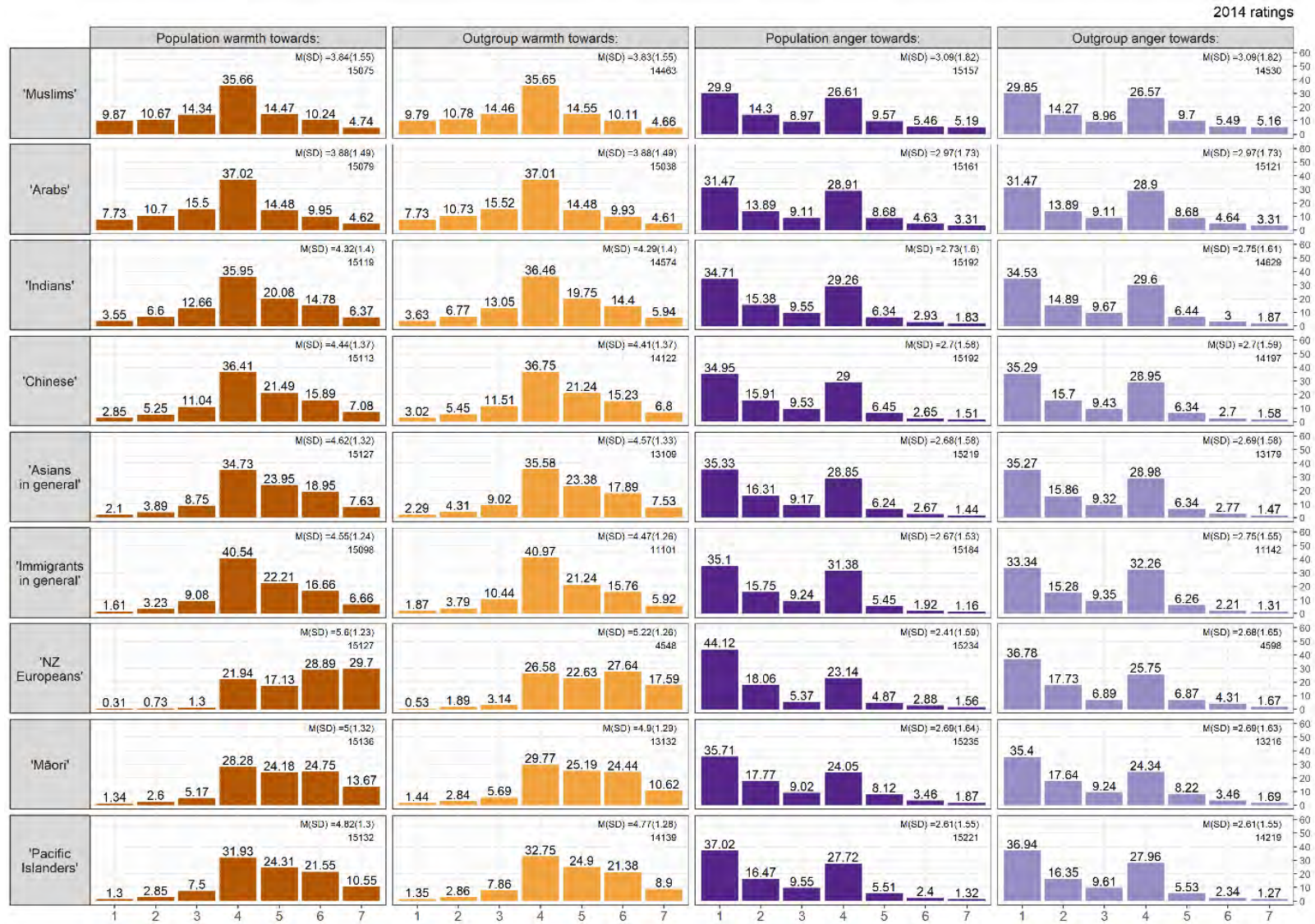


Figure 5. Histograms showing the distribution of thermometer ratings of warmth and anger toward each group in the Time 6 (2014) wave of the NZAVS

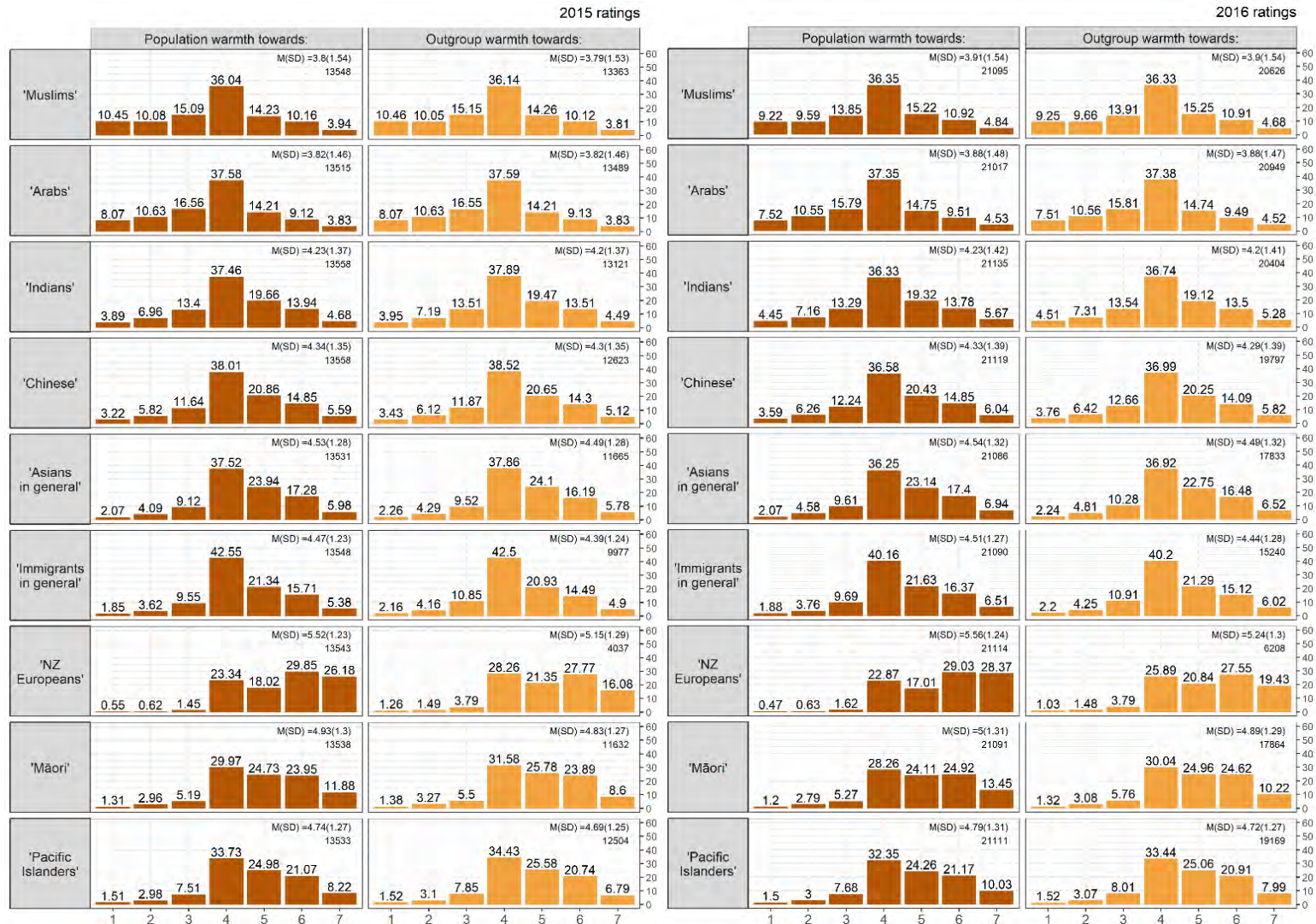


Figure 6. Histograms showing the distribution of thermometer ratings of warmth toward each group in the Time 7 (2015) wave of the NZAVS.

Figure 7. Histograms showing the distribution of thermometer ratings of warmth toward each group in the Time 8 (2016) wave of the NZAVS.

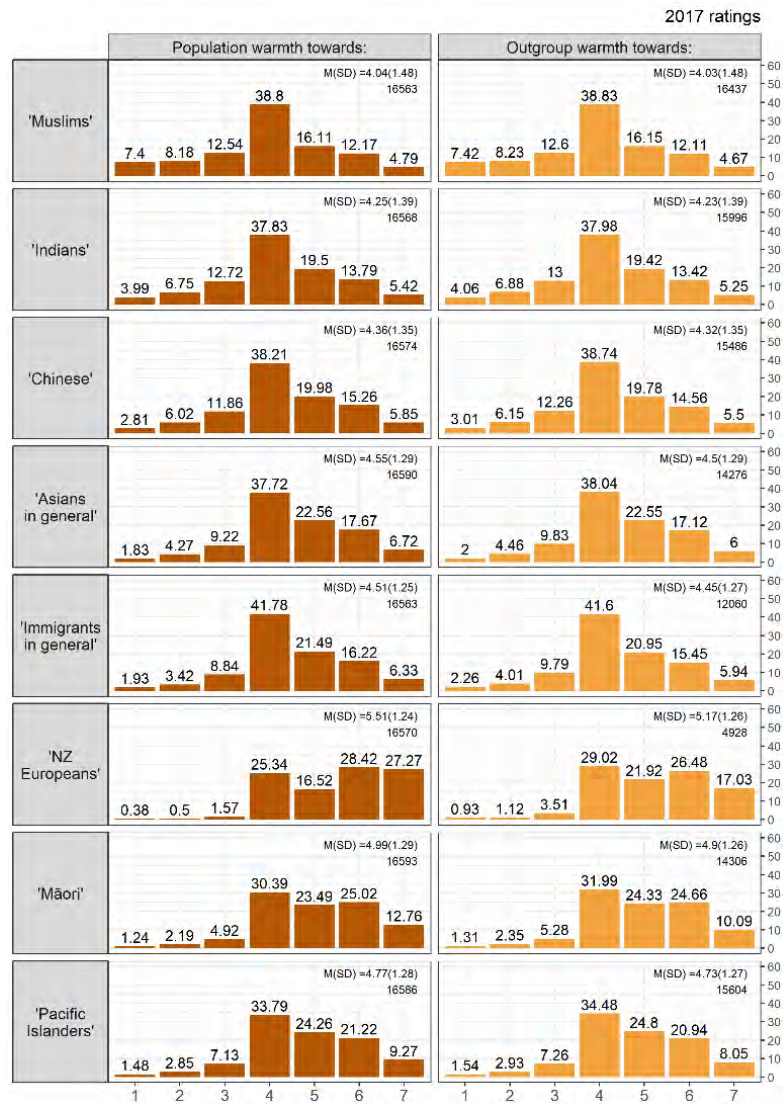


Figure 8. Histograms showing the distribution of thermometer ratings of warmth toward each group in the Time 9 (2017) wave of the NZAVS.

Table 5. Descriptive statistics and correlations for all variables

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.
1. Warmth toward Muslims	-																							
2. Warmth toward Immigrants	.704	-																						
3. Gender (0 women, 1 male)	-.102	-.064	-																					
4. Age	-.095	-.005	.111	-																				
5. Education	.170	.155	-.044	-.198	-																			
6. Household Income (Log)	.062	.068	.064	-.163	.251	-																		
7. Deprivation	-.024	-.047	-.027	-.033	-.142	-.274	-																	
8. Socio-economic Status	.141	.139	-.076	-.086	.559	.295	-.152	-																
9. Māori Ethnicity (0 no, 1 yes)	-.011	-.054	-.031	-.047	-.099	-.061	.157	-.069	-															
10. Pacific Ethnicity (0 no, 1 yes)	.006	.008	-.005	-.064	-.025	-.030	.117	-.022	.105	-														
11. Asian Ethnicity (0 no, 1 yes)	.002	.024	-.015	-.121	.096	-.012	.009	.045	-.046	.004	-													
12. Religious (0 no, 1 yes)	-.025	.038	-.045	.131	-.005	-.064	.038	.002	.020	.082	.057	-												
13. Muslim (0 no, 1 yes)	.044	.009	-.007	-.017	.017	-.007	.018	.010	-.008	.017	.078	.061	-											
14. Parent (0 no, 1 yes)	-.045	-.011	.008	.416	-.110	.044	-.057	-.010	.026	-.025	-.086	.079	-.001	-										
15. Partner (0 no, 1 yes)	-.001	.025	.079	.034	.051	.343	-.189	.098	-.068	-.035	-.022	-.019	-.002	.271	-									
16. Employed (0 no, 1 yes)	.054	.030	.036	-.314	.155	.339	-.067	.123	-.006	-.008	.003	-.055	-.017	-.074	.079	-								
17. Urban (0 rural, 1 urban)	.045	.048	-.007	-.053	.091	.028	.097	.094	-.011	.045	.059	.022	.014	-.085	-.101	-.011	-							
18. Born in NZ (0 no, 1 yes)	-.021	-.077	-.030	-.009	-.158	-.032	.059	-.100	.160	-.023	-.254	-.032	-.033	.007	-.059	.009	-.015	-						
19. Extraversion	.073	.104	-.042	-.011	.026	.098	-.045	.057	.020	.016	-.026	.021	-.005	.070	.072	.052	.003	.001	-					
20. Agreeableness	.203	.214	-.292	.007	.095	.018	-.044	.109	-.053	-.032	-.020	.074	.001	.040	.027	-.010	.018	-.019	.204	-				
21. Conscientiousness	-.014	.026	-.085	.065	-.018	.071	-.059	.020	.002	.022	.003	.042	.010	.076	.074	.007	.004	-.019	.055	.137	-			
22. Neuroticism	-.055	-.100	-.134	-.220	.006	-.053	.043	-.020	.000	.005	.026	-.025	.007	-.119	-.063	.008	.032	.002	-.143	-.035	-.189	-		
23. Openness	.149	.138	.055	-.115	.232	.073	-.040	.144	-.011	-.002	.005	-.057	.004	-.079	-.001	.064	.022	-.067	.188	.234	-.027	-.040	-	
24. Honesty/Humility	.152	.148	-.123	.174	.058	-.030	-.026	.054	-.056	-.045	-.067	.015	-.001	.065	.027	-.084	-.046	-.035	-.063	.208	.099	-.176	.063	-
Mean	4.08	4.52	.37	51.20	5.27	11.37	4.59	54.95	.12	.03	.04	.36	.002	.74	.76	.78	.82	.80	3.87	5.35	5.11	3.44	4.93	5.43
SD	1.47	1.25	.48	13.75	2.77	0.84	2.72	16.13	.32	.16	.20	.48	.045	.44	.43	.42	.39	.40	1.17	0.96	1.02	1.14	1.11	1.17

Note. Correlations above .015 are significant at $p < .05$

Table 6. Multiple regression with demographic predictors for warmth toward 'Muslims', and toward 'Immigrants in general' as a comparison (N = 16,641)

	Warmth toward Muslims						Warmth toward Immigrants					
	<i>b</i>	β	<i>se</i>	<i>t</i>	<i>p</i>	95% <i>CI</i>	<i>b</i>	β	<i>se</i>	<i>t</i>	<i>p</i>	95% <i>CI</i>
Gender (0 women, 1 men)	-.141**	-.046	.025	-5.599	<.001	-.190 -.092	-.044*	-.017	.021	-2.099	.036	-.086 -.003
Age	-.008**	-.076	.001	-8.207	<.001	-.010 -.006	.000	.004	.001	0.441	.659	-.001 .002
Education	.045**	.086	.005	8.931	<.001	.035 .055	.033**	.074	.004	7.759	<.001	.025 .042
Household Income (Log)	.014	.008	.018	0.772	.440	-.022 .050	.029*	.020	.015	1.985	.047	.000 .058
Deprivation	-.001	-.001	.004	-0.164	.870	-.009 .008	-.004	-.010	.004	-1.189	.234	-.012 .003
Socio-economic Index	.004**	.045	.001	4.806	<.001	.002 .006	.003**	.044	.001	4.675	<.001	.002 .005
Māori Ethnicity (0 no, 1 yes)	.044	.010	.035	1.230	.219	-.026 .113	-.082*	-.021	.032	-2.569	.010	-.145 -.020
Pacific Ethnicity (0 no, 1 yes)	.139	.015	.074	1.892	.059	-.005 .283	.148*	.019	.068	2.178	.029	.015 .281
Asian Ethnicity (0 no, 1 yes)	-.067	-.009	.056	-1.187	.235	-.176 .043	.060	.010	.047	1.271	.204	-.033 .153
Religious (0 no, 1 yes)	-.083**	-.027	.024	-3.539	<.001	-.130 -.037	.065*	.025	.020	3.311	.001	.027 .104
Muslim (0 no, 1 yes)	1.413**	.043	.232	6.102	<.001	.959 1.867	.110	.004	.208	0.531	.596	-.297 .518
Parent (0 no, 1 yes)	-.028	-.008	.028	-0.989	.323	-.083 .027	-.060*	-.021	.024	-2.459	.014	-.108 -.012
Partner (0 no, 1 yes)	-.036	-.010	.029	-1.240	.215	-.092 .021	.005	.002	.025	0.203	.840	-.044 .054
Employed (0 no, 1 yes)	.061*	.017	.030	2.026	.043	.002 .120	.038	.013	.026	1.469	.142	-.013 .088
Urban (0 rural, 1 urban)	.117**	.031	.029	4.033	<.001	.060 .174	.116**	.036	.025	4.693	<.001	.067 .164
Born in NZ (0 no, 1 yes)	.011	.003	.028	0.392	.695	-.045 .067	-.134**	-.043	.023	-5.735	<.001	-.180 -.088
Extraversion	.034*	.027	.010	3.308	.001	.014 .054	.056**	.053	.009	6.416	<.001	.039 .073
Agreeableness	.214**	.140	.013	16.029	<.001	.188 .240	.195**	.151	.012	16.884	<.001	.173 .218
Conscientiousness	-.076**	-.053	.011	-6.695	<.001	-.098 -.054	-.030*	-.024	.010	-3.063	.002	-.049 -.011
Neuroticism	-.074**	-.058	.011	-7.052	<.001	-.095 -.054	-.081**	-.075	.009	-9.007	<.001	-.099 -.064
Openness	.082**	.062	.011	7.474	<.001	.061 .104	.059**	.053	.009	6.272	<.001	.041 .078
Honesty-Humility	.151**	.121	.011	14.39	<.001	.131 .172	.105**	.099	.009	11.485	<.001	.087 .123

Note. ** $p < .001$, * $p < .05$

Table 7. Fixed and random effects for the Latent Growth Model of warmth toward ‘Muslims’ between 2012 – 2018.

	<i>b</i> [95% CI]	<i>se</i>	<i>t</i>	<i>p</i>
Warmth toward ‘Muslims’				
<i>Fixed effects</i>				
Intercept	3.684 [3.647, 3.721]	.019	194.453	< .001
Linear slope	.118 [.076, .159]	.021	5.571	< .001
Quadratic slope	-.038 [-.054, -.022]	.008	-4.697	< .001
Cubic slope	.005 [.004, .007]	.001	5.844	< .001
<i>Random effects</i>				
Intercept	1.513 [1.418, 1.609]	.049	31.086	< .001
Linear slope	.256 [.083, .429]	.088	2.903	.004
Quadratic slope	.032 [.005, .059]	.014	2.330	.020
Cubic slope	.000 [.000, .001]	.000	1.966	.049
<i>Covariances</i>				
Intercept – Linear slope	-.030 [-.140, -.079]	.056	-0.545	.586
Intercept – Quadratic slope	-.016 [-.058, .027]	.022	-0.724	.469
Intercept – Cubic slope	.002 [-.003, .007]	.002	0.896	.370
Linear slope – Quadratic slope	-.079 [-.146, -.012]	.034	-2.308	.021
Linear slope – Cubic slope	.007 [-.001, .014]	.004	1.806	.071
Quadratic slope – Cubic slope	-.003 [-.006, .000]	.002	-2.077	.038

Note: *N* = 12,936. Fit indices: loglikelihood = -102318.36, AIC = 204666.71, BIC = 204778.73. Models estimated using Maximum Likelihood with robust estimation of standard errors.

Table 9. Fixed and random effects for the Latent Growth Model of warmth toward ‘Asians in general’ between 2012 – 2018.

	<i>b</i> [95% CI]	<i>se</i>	<i>t</i>	<i>p</i>
Warmth toward ‘Asians in general’				
<i>Fixed effects</i>				
Intercept	4.374 [4.340, 4.408]	.017	252.255	< .001
Linear slope	.233 [.193, .273]	.021	11.300	< .001
Quadratic slope	-.080 [-.096, -.065]	.008	-10.142	< .001
Cubic slope	.008 [.006, .010]	.001	9.077	< .001
<i>Random effects</i>				
Intercept	1.005 [.922, 1.089]	.043	23.569	< .001
Linear slope	.309 [.152, .467]	.080	3.844	< .001
Quadratic slope	.037 [.013, .062]	.012	3.002	.003
Cubic slope	.000 [.000, .001]	.000	2.292	.022
<i>Covariances</i>				
Intercept – Linear slope	-.140 [-.238, -.042]	.050	-2.789	.005
Intercept – Quadratic slope	.030 [-.008, .067]	.019	1.542	.123
Intercept – Cubic slope	-.002 [-.006, .002]	.002	-0.983	.325
Linear slope – Quadratic slope	-.101 [-.162, -.040]	.031	-3.227	.001
Linear slope – Cubic slope	.009 [.002, .015]	.003	2.632	.008
Quadratic slope – Cubic slope	-.004 [-.006, .001]	.001	-2.587	.010

Note: *N* = 12,936. Fit indices: loglikelihood = -97068.81, AIC = 194167.62, BIC = 194279.64. Models estimated using Maximum Likelihood with robust estimation of standard errors.

Table 8. Fixed and random effects for the Latent Growth Model of warmth toward ‘Immigrants in general’ between 2012 – 2018.

	<i>b</i> [95% CI]	<i>se</i>	<i>t</i>	<i>p</i>
Warmth toward ‘immigrants in general’				
<i>Fixed effects</i>				
Intercept	4.445 [4.414, 4.477]	.016	275.427	< .001
Linear slope	.120 [.081, .158]	.020	6.102	< .001
Quadratic slope	-.046 [-.061, -.031]	.008	-6.084	< .001
Cubic slope	.005 [.003, .007]	.001	5.757	< .001
<i>Random effects</i>				
Intercept	0.770 [.695, .846]	.039	19.963	< .001
Linear slope	.191 [.038, .343]	.078	2.446	.014
Quadratic slope	.029 [.005, .053]	.012	2.363	.018
Cubic slope	.000 [.000, .001]	.000	2.276	.023
<i>Covariances</i>				
Intercept – Linear slope	-.026 [-.118, .066]	.047	-0.554	.580
Intercept – Quadratic slope	-.002 [-.039, .034]	.018	-0.127	.899
Intercept – Cubic slope	.001 [-.003, .005]	.002	0.410	.682
Linear slope – Quadratic slope	-.068 [-.128, -.008]	.031	-2.209	.027
Linear slope – Cubic slope	.006 [.000, .013]	.003	1.943	.052
Quadratic slope – Cubic slope	-.003 [-.006, .000]	.001	-2.268	.023

Note: *N* = 12,936. Fit indices: loglikelihood = -94536.98, AIC = 189103.95, BIC = 189215.97. Models estimated using Maximum Likelihood with robust estimation of standard errors.

Table 10. Fixed and random effects for the Latent Growth Model of warmth toward ‘Chinese’ between 2012 – 2018.

	<i>b</i> [95% CI]	<i>se</i>	<i>t</i>	<i>p</i>
Warmth toward ‘Chinese’				
<i>Fixed effects</i>				
Intercept	4.191 [4.167, 4.224]	.017	244.998	< .001
Linear slope	.249 [.210, .288]	.020	12.389	< .001
Quadratic slope	-.088 [-.103, -.073]	.008	-11.308	< .001
Cubic slope	.009 [.007, .011]	.001	10.322	< .001
<i>Random effects</i>				
Intercept	0.999 [.918, 1.079]	.041	24.354	< .001
Linear slope	.168 [.055, .280]	.057	2.916	.004
Quadratic slope	.022 [.007, .038]	.008	2.781	.005
Cubic slope	.000 [.000, .000]	.000	2.181	.029
<i>Covariances</i>				
Intercept – Linear slope	-.024 [-.107, .059]	.042	-0.572	.568
Intercept – Quadratic slope	-.004 [-.035, .027]	.016	-0.251	.801
Intercept – Cubic slope	.001 [-.002, .004]	.002	0.560	.575
Linear slope – Quadratic slope	-.053 [-.096, -.012]	.021	-2.555	.011
Linear slope – Cubic slope	.004 [.000, .009]	.002	1.952	.051
Quadratic slope – Cubic slope	-.002 [-.004, .000]	.001	-2.413	.016

Note: *N* = 12,936. Fit indices: loglikelihood = -97530.35, AIC = 195090.71, BIC = 195202.72. Models estimated using Maximum Likelihood with robust estimation of standard errors.

Table 11. Fixed and random effects for the Latent Growth Model of warmth toward 'Indians' between 2012 – 2018.

	<i>b</i> [95% CI]	<i>se</i>	<i>t</i>	<i>p</i>
Warmth toward 'Indians'				
<i>Fixed effects</i>				
Intercept	4.103 [4.068, 4.138]	.018	231.693	< .001
Linear slope	.220 [.180, .260]	.021	10.690	< .001
Quadratic slope	-.075 [-.090, -.059]	.008	-9.417	< .001
Cubic slope	.007 [.006, .009]	.001	8.484	< .001
<i>Random effects</i>				
Intercept	1.172 [1.083, 1.261]	.045	25.799	< .001
Linear slope	.277 [.116, .438]	.082	3.375	.001
Quadratic slope	.037 [.012, .062]	.013	2.913	.004
Cubic slope	.000 [.000, .001]	.000	2.533	.011
<i>Covariances</i>				
Intercept – Linear slope	-.096 [-.198, .007]	.052	-1.831	.067
Intercept – Quadratic slope	.015 [-.025, .054]	.020	0.732	.464
Intercept – Cubic slope	-.001 [-.005, .004]	.002	-0.317	.751
Linear slope – Quadratic slope	-.092 [-.155, -.030]	.032	-2.884	.004
Linear slope – Cubic slope	.008 [.002, .015]	.003	2.398	.016
Quadratic slope – Cubic slope	-.004 [-.006, -.001]	.001	-2.653	.008

Note: *N* = 12,936. Fit indices: loglikelihood = -98644.52, AIC = 197319.04, BIC = 197431.05. Model estimated using Maximum Likelihood with robust estimation of standard errors.

The results of the models assessing each group are displayed in Tables 7-11, with the fixed effect coefficients indicating the extent and pattern of change in warmth. Looking at warmth toward Muslims, the cubic slope ($b = .005$, $p < .001$), quadratic slope ($b = -.038$, $p < .001$), and linear slope ($b = .118$, $p < .001$) were each significant, indicating non-linear change in warmth over time. Similarly, the models assessing warmth toward immigrants, Asian peoples, Chinese people, and Indians also showed significant linear, quadratic, and cubic effects.

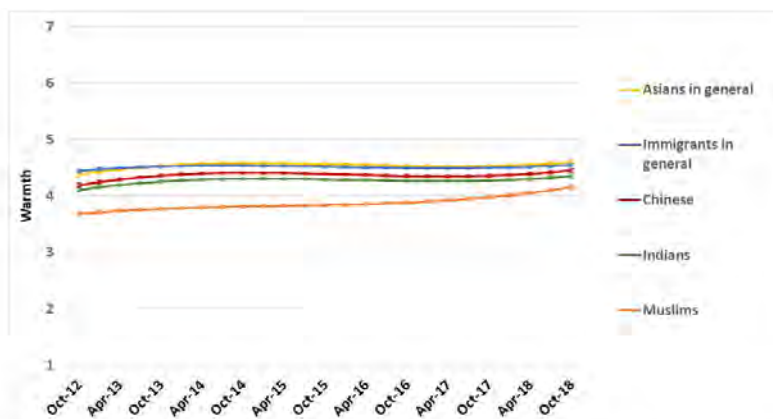


Figure 9. Latent growth curves representing the average rate of change in feeling thermometer ratings toward 'Asians in general', 'Immigrants in general', 'Chinese', 'Indians' and 'Muslims.' Growth curves were estimated separately for ratings of each group, and the curves represent model-implied values based on the parameters reported in Tables 7-11.

The trajectory of change in warmth toward Muslims (alongside the comparative groups) is displayed in Figure 9. The figure plots model-implied levels of warmth estimated at 3-monthly intervals between October 2012 and October 2018. As shown in Figure 9, warmth toward Muslims (3.68) was lower than the mid-point of the scale in October 2012, and the lowest of any group examined. However, warmth increased (at varying rates) over the six-year period, and was notably higher (4.16; although still at the neutral point of the scale) by October 2018.

Whereas the trajectory of change in warmth toward the other groups examined exhibited periods of slight decline in warmth, the trajectory of change in warmth toward Muslims was always positive. Overall, between October 2012 to October 2018, warmth toward Muslims rose by .474 (on a scale from 1-7), which was the largest increase in warmth toward any group examined.

DISCUSSION

In this paper, we present a summary of feeling thermometer data on warmth toward Muslims from the New Zealand Attitudes and Values Study (NZAVS). We present data from the Time 4 (2012) to Time 9 (2017) waves of the NZAVS, as we included warmth toward Muslims for the first time in 2012.

Summary of Findings

Our results are organized in three sections: (a) a comprehensive description of the levels of warmth and anger toward Muslims and other groups in New Zealand, and the distribution of these ratings, (b) a regression model summarizing the extent to which a broad range of demographics and aspects of personality are associated with low versus high levels of warmth toward Muslims, and (c) a latent growth curve model showing how feelings toward Muslims changed from 2012 onwards. We summarize the results of each section below.

The histograms summarizing feeling thermometer ratings (Figures 3-8) indicate that, across the years, warmth ratings toward Muslims were relatively normally distributed, yet lower on average than warmth ratings toward all the other groups we examined. This was apparent in analyses of both the overall sample for each wave, and the outgroup estimate excluding members of each group when estimating warmth ratings toward that group. From 2012 to 2017, there was consistently less warmth expressed toward Muslims than other groups in New Zealand. The same pattern held for the years in which we also measured ratings of 'Arabs.'

Another trend that can be seen in these histograms is that, across years, there are consistently fewer people in the population rating below the midpoint (i.e., ratings below neutral) in feelings of warmth toward NZ Europeans, Pacific peoples, and Māori. This pattern remains consistent when looking specifically at outgroup ratings (e.g., warmth toward NZ Europeans excluding the

ratings of NZ European participants). For ratings of NZ Europeans, Pacific people and Māori, the effective thermometer scale looks more like it is truncated and effectively ranges from 'neutral' to 'most warm' rather than from 'least warm' through 'neutral' to 'most warm.'

The histograms for anger show a different pattern of distributions to those measuring feelings of warmth. A sizeable proportion of the NZ population (between 30-40%) expressed no anger toward any of the groups, including Muslims. However, even here we see variation. As can be seen in Figures 3-8, across the years in which we measured anger, consistently more people rated no anger toward all other groups than the amount of people who rated no anger toward Muslims or Arabs (the exact proportions are reported in Tables 2-4). As with low warmth, more people in New Zealand report at least a little anger toward Muslims and Arabs relative to the other groups we measured feelings toward.

The regression model presented in the second section of results provides a profile of the demographics and aspects of personality that are associated with lower or higher levels of concurrent warmth toward Muslims. This model assesses the correlation of each demographic or personality trait with warmth toward Muslims, when statistically adjusting for all the other factors in the model. For example, it documents the association between age and warmth toward Muslims adjusting for gender, education, Extraversion and so forth. Another way to think about this regression model is that it provides a profile of the factors associated with low or high warmth toward Muslims, or a demographic and personality segmentation of those factors.

The *b* values in the model represent the unstandardized regression parameters. These parameters are extremely useful because they provide the actual predicted units of change in warmth associated with a one-unit change in the predictor, adjusting for all other predictors in the model. Put another way, for each one-unit increase in a predictor variable, there is, on average, a corresponding *b* unit change in the level of warmth toward Muslims. The β parameters in the model provide much the same information, but in standard deviation units. This is useful because it allows one to get a sense of the relative magnitude of each predictor compared to the others, as they are on a common (standard deviation unit) metric.

Walking through the regression parameters reported in Table 6, men tended, on average, to express less warmth toward Muslims than women (specifically, they tended on average to report .141 units less warmth). Older people were also less warm toward Muslims, again on average. Those with a higher level of education were warmer toward Muslims, as were those with a higher socio-economic index, and those living in urban areas. Differences in the regional deprivation of one's neighbourhood, household income, whether one was a parent, had a partner, were born in New Zealand, and one's own ethnicity were all unassociated with levels of warmth toward Muslims. Unsurprisingly, Muslims expressed a high level of warmth toward Muslims.

The regression model reported in Table 6 also indicates that religiously affiliated people were less warm toward Muslims than were non-religiously affiliated people. This observation is qualified by the in-depth analyses of NZAVS data conducted by Shaver et al. (2016). Shaver et al. showed that weakly religiously identified people exhibited less warmth to Muslims than non-affiliated people, and that among the religiously affiliated, higher warmth to Muslims was observed only among highly religiously-identified, church-attending religious affiliates (see also Highland et al., 2019).

With regard to personality, higher levels of Extraversion, Agreeableness, Openness to Experience and Honesty-Humility were associated with higher warmth toward Muslims. Conversely, higher levels of Conscientiousness and Neuroticism (or low Emotional Stability) were associated with lower warmth toward Muslims. These personality results are broadly consistent with the links between personality and prejudice in other domains (see Sibley & Duckitt, 2008, for a meta-analysis and review), and were included in this model simply to statistically adjust for their associations when examining the correlations between demographic factors and thermometer ratings.

Although many of the demographics in the regression model were significant, it is critical to recognize that many of the associations in the model were extremely small in magnitude. They describe, at best, extremely weak effects. One can get a sense of this by looking at the standardized coefficients in the model, which were all below .30 (a correlation value of .30 is typically described as a moderate effect, and one of .50 as strong). Of the various demographics included in the regression model, education showed the largest effect size (although still being quite a small effect), with $\beta = .086$. This is promising as levels of education are something that a society can change (also see Osborne, Satherley, Yogeeswaran, Hawi, & Sibley, 2019).

We hope that this regression model will be of use for people seeking to understand the segmentation and prediction of anti-Muslim sentiment in the population. We hope that this model will also provide an evidence base moving forward when developing prejudice reduction interventions tailored toward specific segments of the population.

The third section of the results describes a series of Latent Growth Curves assessing the level and rate of change in feeling thermometer ratings of warmth toward Muslims from October 2012 to October 2018. We also report separate growth models examining change in warmth ratings toward 'Asians in general', 'Immigrants in general', 'Chinese', and 'Indians' as a comparison. We chose to include ratings of these groups as a comparison because they showed the most similar levels and distributions to rating of Muslims in our histograms.⁹

The results of our growth curve analyses indicate that ratings of warmth toward all five groups in our analysis increased from October 2012 to October 2018. The slopes reported in Figure 9 further indicate that change may not

⁹ We did not estimate a growth model for thermometer ratings of 'Arabs' because warmth toward them was not included in the Time 9 questionnaire (it was swapped out to measure warmth toward

refugees instead). We also did not estimate growth curves for anger ratings as they likewise had been included at a more limited set of study waves.

be entirely linear, perhaps plateauing for some period of time, and accelerating at other times.¹⁰ As can also be seen in Figure 9, overall levels of warmth toward Muslims remained lower than warmth ratings of ‘Asians in general’, ‘Immigrants in general’, ‘Chinese’ and ‘Indians’ throughout the October 2012 to October 2018. Moreover, while warmth toward all groups increased over this time period, warmth toward Muslims showed the greatest increase over time. There is most certainly further room for improvement, but we are tracking an upward trend in warmth.¹¹

Ruling Out Response Bias in Thermometer Ratings

Readers might reasonably wonder if people respond honestly to feeling thermometer ratings. Perhaps no one is willing to rate that they feel little or no warmth toward other groups, or perhaps our scales underestimate the number of people who actually feel this way? For a start, we do see considerable variation in feeling thermometer ratings toward a range of different groups, with people responding using the full range of the scale for a number of groups. Our results certainly indicate that some people rate that they feel little or no warmth toward Muslims. This does not, however, speak to the possibility that low levels of warmth may be underestimated in survey studies (they would be underestimated if perhaps some people were unwilling to admit the magnitude of their consciously-held low feelings of warmth, and this would inflate our estimate of the true score).

One way to see if affect thermometer ratings are biased because some people are unwilling to admit their consciously-held low feelings of warmth is to check whether warmth ratings are systematically correlated with lie detection items (known as socially desirable responding). The Time 1 (2009) NZAVS questionnaire included two marker items from Paulhus’ (1991) inventory of socially desirable responding. One item assessed impression management (i.e., I ... ‘Don’t care to know what other people really think of me’) and one item assessed self-deceptive enhancement (i.e., I ... ‘Don’t gossip about other people’s business’). These are the two primary factors underlying response bias due to socially desirable responding.

The idea behind including these items is that very few people should be able to honestly agree with them. Most

of us care at least a little regarding what others think about us, and most of us gossip at least a little. One can then check to see if the tendency to agree with these items (i.e., reporting that one does not care what others think or ever gossips) correlates with other self-report measures of interest. To the extent to which they reliably correlate with other scales, those scales may be contaminated by response bias due to impression management and self-deceptive enhancement. This contamination may then be adjusted for statistically to obtain a more reliable estimate of the true level of warmth or anger.

Fortunately, analysis of the Time 1 NZAVS ($N = 6,518$) indicated that scores on the two social desirability marker items were not, or at best extremely weakly, correlated with the various affect thermometer ratings included in the questionnaire. This is good news as it indicates little-to-no response bias due to socially desirable responding. For each feeling thermometer ratings, the correlation with the impression management and self-deceptive enhancement items were, respectively: Māori ($r_s = .023, .025$), Pacific peoples ($r_s = -.009, .035$), Asians ($r_s = -.015, .054$), NZ Europeans ($r_s = -.042, -.006$), Chinese ($r_s = -.014, .049$), immigrants in general ($r_s = -.031, .064$), overweight people ($r_s = .027, .031$), and Americans ($r_s = -.012, .042$).

It is also possible that the specific thermometer items included in the NZAVS may be biased by a comparison effect, where people are subjectively rating their warmth toward each group relative to the list of other groups also included in the scale. This is a difficult criticism to entirely rule out, and there is research suggesting that people are influenced by previous Likert items more generally when completing questionnaires. For instance, it is generally recommended that more general questions should be asked first and specific questions afterwards, as asking specific questions can bias responses to more general questions in the same domain (see for example, Garbarski, Schaeffer & Dykema, 2015). If there is a bias introduced by some sort of comparison effect in the thermometer ratings included in the NZAVS, then the point that people feel the least warmth toward Muslims of all the groups we measure still stands. Unfortunately, we do not have experimental data comparing warmth ratings of groups when the order or composition of the target groups included in the list differs across conditions.

¹⁰ Although the data do indicate subtle cubic effects in the rate of change, it is risky to read too much into possible curvilinear patterns without more waves of data. We are, quite literally, fitting curves with only six annual time points, or six points of observation per person (and less in some cases, as not all participants completed all six waves). Our Monte Carlo simulations indicate that this provides adequate statistical power to estimate overall rates of change given the sample size and pattern of missingness in the NZAVS dataset (see Sibley & Milojev, 2014). However, it is by no means enough to test subtler trends, such as possible cyclical trends or oscillations across election cycles or changes in government. We also do not have enough waves of data (yet) to test piecewise growth models, where one rate of change is estimated prior to some point in time (say a major world event) and another rate of change is estimated following it. In sum, we would not make too much of the slight non-linear curves in these growth models - perhaps they are patterns of change as we emerge from the Global Financial Crisis? Or perhaps New Zealander’s attitudes have become more tolerant in recent years following the rise in populism seen in some other countries? Both are possibilities, although highly post-hoc,

and somewhat tenuous explanations. Yet a third possibility is the increased number of migrants to New Zealand overall during this period. As our migrant population increases, in particular urban areas and populations attitudes to migrants may have warmed; conversely, in areas with lower direct experience of migrants, attitudes may not have warmed to the same level. Regardless, this should not overshadow the primary result: rates of warmth have increased.

¹¹ Some readers may wonder if the rate of increase in warmth toward Muslims might simply be explained by regression to the mean. The argument here is that, statistically, observations that are at the extreme, i.e., those most above or below the mean, are the most likely to move toward the mean over time. We think that this is unlikely because if regression to the mean were driving this trend then one might expect to see a corresponding decrease in warmth toward Asians (as those ratings are approximately as much above the midpoint as ratings of Muslims are below it). The trends are also consistent across six annual waves of data, rather than say, just two.

The NZAVS also included a warmth rating of ‘Arabs’ from Time 4-Time 8, but this was swapped out to include ratings toward ‘Refugees’ in Time 9. Warmth toward Muslims and warmth towards Arabs were highly correlated at each time point (e.g., at Time 4, $r = .834$; at Time 5, $r = .823$; at Time 6, $r = .845$; at Time 7, $r = .847$; at Time 8, $r = .842$), suggesting that responses to these two measures are relatively indistinguishable in the New Zealand context. Of course, this coupling of attitudes toward Muslims and Arabs is context specific and would likely differ in contexts where people were more exposed, or at least aware of, Muslim communities in different parts of the world and of different ethnic groups, and also people from the Middle East who affiliate with religions other than Islam. In New Zealand, however, the concepts of Muslim and Arab are strongly linked as one overall concept.

Concluding Comments

In the wake of the March 15th Christchurch terrorist attacks, our research group feels that it is important to provide aggregate or overall summary data on levels of prejudice toward Muslims in New Zealand. It is our hope that such information will help policy-makers, community groups, and other bodies reflect on and alleviate prejudices faced by Muslims in New Zealand. Rather than presenting these data piecemeal, or providing some specific findings to some groups, but different analyses to others, we opted to present, as comprehensively as we could, a summary of what we

know about levels of warmth and anger toward Muslims and other groups in New Zealand. As the curators of the NZAVS, we think that these findings should be in the public domain, accessible to all.

There is a lot of subjectivity in how one can present research findings, particularly in the social sciences. We have deliberately stuck close to these data, and tried to present findings from the NZAVS in a broadly descriptive way so that they may be understood by a wide audience while keeping our (subjective) interpretation of the findings fairly brief. We wanted to let our findings speak for themselves. We think that they do.

We hope that, in some small way, our findings may be of benefit to the community moving forward, to Muslim members of our community in particular, to government departments and other community groups seeking to understand the level of prejudice in New Zealand. We hope that our findings will be useful for those aiming to design policy or evaluate interventions to increase tolerance and challenge prejudice and anti-Islamic attitudes, such as the Human Rights Commission and Ministry of Social Development. Finally, we hope that our findings will be useful for those contributing their voice to the inevitable and enduring conversation that we, as a bicultural nation but also a multicultural nation, must have about what our national identity is, and the values and beliefs that we think should represent us (and those that it seems sometimes do whether we like it or not).

References

- Atkinson, J., Salmond, C., & Crampton, P. (2014). NZDep 2013 Index of Deprivation. University of Otago.
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11, 150-166.
- Barlow, F. K., Hornsey, M., Hayward, L., Houkamau, C. A., Kang, J., Milojev, P., & Sibley, C. G. (2019). Why do we hold mixed emotions about racial outgroups? A case for affect matching. *Psychological Science*, 30, 917-929.
- Bergh, R., Akrami, N., Sidanius, J., & Sibley, C. G. (2016). Is group membership necessary for understanding generalized prejudice? A re-evaluation of why prejudices are interrelated. *Journal of Personality and Social Psychology*, 111, 367-395.
- Duckitt, J. (2001). A dual-process cognitive-motivational theory of ideology and prejudice. In M. P. Zanna, (Ed.), *Advances in experimental social psychology*, (Vol. 33, pp. 41-113). New York: Academic Press.
- Duckitt, J., & Sibley, C. G. (2007). Right-Wing Authoritarianism, Social Dominance Orientation and the dimensions of generalized prejudice. *European Journal of Personality*, 21, 113-130.
- Garbarski, D., Schaeffer, N. C., & Dykema, J. (2015). The effects of response option order and question order on self-rated health. *Quality of Life Research*, 24(6), 1443-1453.
- Greaves, L. M., Rasheed, A., D’Souza, S., Shackleton, N., Oldfield, L. D., Sibley, C. G., ... & Bulbulia, J. (2020). Comparative study of attitudes to religious groups in New Zealand reveals Muslim-specific prejudice. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 1-20.
- Hamley, L., Houkamau, C. A., Osborne, D., Barlow, F. K., & Sibley, C. G. (2020). Ingroup love or outgroup hate (or both)? Mapping distinct bias profiles in the population. *Personality and Social Psychology Bulletin*, 46, 171-188.
- Hawi, D., Osborne, D., Bulbulia, J., & Sibley, C. G. (2019). Terrorism anxiety and attitudes toward Muslims. *New Zealand Journal of Psychology*, 48, 80-89.
- Highland, B. R., Troughton, G., Shaver, J., Barrett, J. L., Sibley, C. G., & Bulbulia, J. (2019). Attitudes to religion predict warmth for Muslims in New Zealand. *New Zealand Journal of Psychology*, 48, 122-132.
- Huang, Y., Greaves, L. M., & Sibley, C. G., (2014). Archive of additional (non-questionnaire) materials posted to participants. *NZAVS Technical Documents*, e07.
- Kabir, S. N., & Bourk, M. (2012). Representing Islam and Muslims in New Zealand newspapers. *Journal of Muslim Minority Affairs*, 32, 324-338.
- Kinder, D. R., & Drake, K. W. (2009). Myrdal’s prediction. *Political Psychology*, 30, 539-568.
- Milne, B. J., Byun, U., & Lee, A. (2013). *New Zealand Socio-economic Index 2006*.
- Milojev, P., Osborne, D., Greaves, L. M., Barlow, F. K., & Sibley, C. G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. *Journal of Research in Personality*, 47, 936-944.
- Osborne, D., & Sibley, C. G. (2017). Identifying ‘types’ of ideologies and intergroup biases: advancing a person-centred approach to social psychology. *European Review of Social Psychology*, 28, 288-332.
- Osborne, D., Satherley, N., Yogeeswaran, K., Hawi, D., & Sibley, C. (2019). White nationalism and multiculturalism support: Investigating the interactive effects of White identity and national attachment on support for

- multiculturalism. *New Zealand Journal of Psychology*, 48(1), 62-70.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes*. San Diego: Academic Press.
- Perry, R., Priest, N., Paradies, Y., Barlow, F. K., & Sibley, C. G. (2018). Barriers to multiculturalism: ingroup favoritism and outgroup hostility are independently associated with policy opposition. *Social Psychology and Personality Science*, 9, 89-98.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751.
- Rubin, D. B. (1978). Multiple imputations in sample surveys—a phenomenological Bayesian approach to nonresponse. In *Proceedings of the survey research methods section of the American Statistical Association* (Vol. 1, pp. 20-34). American Statistical Association.
- Salmond, C., Crampton, P., & Atkinson, J. (2007). *NZDep2006 Index of Deprivation*. Wellington: Department of Public Health; Atkinson, J., Salmond, C., & Crampton, P. (2014). *NZDep2013 Index of Deprivation*. University of Otago.
- Satherley, N., Milojev, P., Greaves, L. M., Huang, Y., Osborne, D., Bulbulia, J., & Sibley, C. G. (2015). Demographic and psychological predictors of panel attrition: evidence from the New Zealand Attitudes and Values Study. *PLoS ONE*, 10, e0121950.
- Satherley, N., & Sibley, C. G. (2018). The Modern Racism toward Māori Scale. *New Zealand Journal of Psychology*, 47, 4-12.
- Sengupta, N. K., Barlow, F. K., & Sibley, C. G. (2012). Intergroup contact and post-colonial ideology: outgroup contact ameliorates symbolic exclusion but not historical negation. *International Journal of Intercultural Relations*, 36, 506-517.
- Shaver, J., Sibley, C. G., Osborne, D., Bulbulia, J. (2017). News exposure predicts anti-Muslim prejudice. *PLoS ONE*, 12, e0174606.
- Shaver, J., Troughton, G., Sibley, C. G., Bulbulia, J. (2016). Religion and the unmaking of prejudice towards Muslims: evidence from a large national sample. *PLoS ONE*, 11, e0150209.
- Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge: Cambridge University Press.
- Sibley, C. G. (2012). The Mini-IPIP6: Item Response Theory analysis of a short measure of the big-six factors of personality in New Zealand. *New Zealand Journal of Psychology*, 41, 21-31.
- Sibley, C. G. (2019a) Sampling procedure and sample details for the New Zealand Attitudes and Values Study. NZAVS Technical Documents, e01. Available at: www.nzavs.auckland.ac.nz
- Sibley, C. G. (2019b) NZAVS Data Dictionary. NZAVS Technical Documents, e02 Available at: www.nzavs.auckland.ac.nz
- Sibley, C. G. (2019c). Archive of NZAVS questionnaires. *NZAVS Technical Documents*, e06. Available at: www.nzavs.auckland.ac.nz
- Sibley, C. G. (2019d). Procedures for estimating post-stratification NZAVS sample weights. *NZAVS Technical Documents*, e08. Available at: www.nzavs.auckland.ac.nz
- Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review*, 12, 248-279.
- Sibley, C. G., Luyten, N., Purnomo, M., Moberly, A., Wootton, L. W., Hammond, M. D., Sengupta, N., Perry, R., West-Newman, T., Wilson, M. S., McLellan, L., Hoverd, W. J., & Robertson, A. (2011). The Mini-IPIP6: Validation and extension of a short measure of the Big-Six factors of personality in New Zealand. *New Zealand Journal of Psychology*, 40, 142-159.
- Sibley, C. G., & Milojev, P. (2014). Power estimation of slope growth factors in the NZAVS using Monte Carlo simulation. NZAVS Technical Documents, e19. Available at: www.nzavs.auckland.ac.nz
- Sibley, C. G., & Pirie, D. J. (2013). Personality in New Zealand: Scale norms and demographic differences in the Mini-IPIP6. *New Zealand Journal of Psychology*, 42, 13-30.
- Sibley, C. G., & Ward, C. (2013). Measuring the preconditions for a successful multicultural society: a barometer test of New Zealand. *International Journal of Intercultural Relations*, 37, 700-713.
- Statistics New Zealand (2014). *Quickstats about great Christchurch*. Statistics New Zealand.
- Teaiwa, T., & Mallon, S. (2005). Ambivalent kinships? Pacific people in New Zealand. In J. H. Liu, T. McCreanor, T. McIntosh, & T. Teaiwa (Eds.), *New Zealand identities: Departures and destinations* (pp. 207–229). Wellington, New Zealand: Victoria University Press.
- Wilson, M. S. (2019). A community-based test of the Dual Process Model of Intergroup Relations: Predicting attitudes towards Christians, Muslims, Hindus, Jews, and Atheists. *New Zealand Journal of Psychology (Online)*, 48(1), 133-139.
- Yogeeswaran, K., Afzali, M. U., Andrews, N. P., Chivers, E. A., Wang, M-J., Devos, T., Sibley, C. G. (2019). Exploring New Zealand national identity and its importance for attitudes toward Muslims and diversity. *New Zealand Journal of Psychology*, 48, 29-35.

Corresponding Author

Chris Sibley
 School of Psychology, University of Auckland,
 E-mail: c.sibley@auckland.ac.nz

Acknowledgements: This research was supported by a grant from the Templeton Religion Trust (TRT0196). The funders have no role in NZAVS study design, data collection and analysis, decision to publish, or preparation of scientific reports or manuscripts for publication using NZAVS data. Syntax for the models reported in this paper, supplementary materials providing full information for the distribution of all feeling thermometer scores, and high resolution images of Figures 3-9 for use in presentations, are available on the bibliography page at the NZAVS website: www.nzavs.auckland.ac.nz

Appendix: NZAVS Sampling Procedure

Sampling Procedure -- NZAVS Time 4 (2012)

The Time 4 (2012) NZAVS contained responses from 12,179 participants (6,807 retained from one or more previous wave, 5,107 new additions from booster sampling, and 265 unmatched participants or unsolicited opt-ins). Informal analysis indicates that unsolicited opt-ins were often the partners of existing participants. The sample retained 4,053 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 62.2% over three years). The sample retained 5,762 participants from the full Time 3 (2011) sample (a retention rate of 83.7% from the previous year). Participants were posted a copy of the questionnaire, with a second postal follow-up two months later. Participants who provided an email address were also emailed and invited to complete an online version if they preferred.

Non-respondents were emailed a follow-up reminder email approximately two months later. Three attempts were then made using each provided phone number (typically home and cell phone) to contact non-respondents to encourage participation. These attempts were made on separate days, approximately one week apart. When possible, a phone message was left for each phone number after the third attempt. Participants were also posted a pamphlet outlining recent findings from the study mid-way through the year. Finally, participants were posted a Season's Greetings from the NZAVS research team, and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw for a total pool of \$NZ 1,000). Participants were informed that the draw would happen automatically and winners contacted. The Season's Greetings card also asked participants to contact us (online, email or phone) to let us know if any of their contact details had changed before the prize draw was conducted. These additional materials are presented by Huang, Greaves, and Sibley (2014) in an online NZAVS technical report.

To boost sample size at Time 4 and increase sample diversity for subsequent waves, five independent booster samples using different sample frames were also conducted. Booster sampling was conducted without replacement (i.e., all people included in previous sample frames were identified and removed from the electoral roll before generation of the new sample frames). The first sample frame consisted of a randomly selected sample of 20,000 people from the 2012 New Zealand Electoral, and who were currently residing in New Zealand (one can be registered to vote in New Zealand but living overseas). A total of 2,429 participants responded to this booster sample (response rate = 12.33% when adjusting for the 98.5% accuracy of the 2012 electoral roll). The second sample frame consisted of a regional booster of 10,000 people randomly selected from people listed in the 2012 Electoral Roll who lived in the Auckland region. A total of 890 participants responded to this booster sample (adjusted response rate = 9.04). The Auckland region was oversampled because it is the fastest growing and most ethnically diverse region of the country with an increasing number of Asian

and Pacific peoples in particular. The questionnaire used for this Auckland sample was longer than the standard NZAVS questionnaire, and contained additional unrelated questions that are not included in the NZAVS dataset (these related to the use of community facilities). Exit interviews conducted during Time 5 indicated that the longer length of this questionnaire may have contributed to the low response rate in this case.

The third sample frame consisted of 3,000 people randomly selected from the 2012 Electoral Roll who lived in the Christchurch region. A total of 332 participants responded to this booster sample (adjusted response rate = 11.24%). The Christchurch region was oversampled because it has experienced significant hardship and change due to the Christchurch earthquakes of 2010 and 2011 with many people moving out of the region (Statistics New Zealand, 2013) and problems with mail delivery with some city zones being placed under restricted entry due to safety concerns and considerable infrastructure destroyed).

The fourth sample frame consisted of 9,000 respondents selected from meshblock area units across the country that were moderate-to-high in deprivation according to the index developed by Salmond, Crampton and Atkinson (2007). Regions with levels of deprivation were selected using scores on the decile-ranked NZ Deprivation index from 6-10, with 10 being the most deprived). This sample frame used scaled weighting so that people in increasingly deprived regions were increasingly more likely to be selected (with random sampling of people within regions that had a given level of deprivation). The scaling factor was as follows: $n_i = n_{base} * weight_i$, where $n_{base} = 600$, and $weight_i$ ranged from 1 to 5 and increased by 1 for each one-unit increase in deprivation score. Thus, 600 people were randomly selected from regions with a deprivation score of 6, 1,200 people were randomly selected from regions with a deprivation score of 7, and so on. This sampling strategy was designed to increase the representativeness of the sample across regions with different levels of deprivation, as the NZAVS showed increased attrition in increasingly more deprived regions over the first three years of the study. A total of 767 participants responded to this booster sample (adjusted response rate = 8.65%). The fifth sample frame consisted of 9,000 people randomly selected from those who indicated on the 2012 Electoral Roll that they were of Māori ethnicity (ethnic affiliation as Māori is listed on the roll, but other ethnic affiliations are not). A total of 689 participants responded to this booster sample (adjusted response rate = 7.78%). The questionnaire administered to the Māori booster sample included questions specifically designed for Māori.

Sampling Procedure -- NZAVS Time 5 (2013)

The Time 5 (2013) NZAVS contained responses from 18,261 participants (10,502 retained from one or more previous wave, 7,579 new additions from booster sampling, and 179 unmatched participants or unsolicited opt-ins). Informal analysis indicates that unsolicited opt-ins were often the partners of existing participants. The sample retained 3,934 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate

of 60.4% over four years). The sample retained 9,844 participants from the full Time 4 (2011) sample (a retention rate of 80.8% from the previous year). Participants were posted a copy of the questionnaire, with a second postal follow-up two months later. Participants who provided an email address were also emailed and invited to complete an online version if they preferred. As described in the Time 4 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw. We also posted our yearly pamphlet summarizing key research findings published during the current wave of the study.

To boost sample size and increase sample diversity for subsequent waves, two booster samples were also conducted by selecting people from the New Zealand electoral roll. As with previous booster samples, sampling was conducted without replacement (i.e., all people included in previous sample frames were identified and removed from the 2014 roll). The first sample frame consisted of 70,000 people aged from 18-60 randomly selected from the 2014 New Zealand Electoral Roll, who were currently residing in New Zealand (one can be registered to vote in New Zealand but living overseas). The New Zealand Electoral Roll contains participants' date of birth (within a one-year window), and we limited our frame to people who 60 or younger, due to our aim of retaining participants for the following 15 years. A total of 7487 participants responded to this booster sample (response rate = 10.9% when adjusting for the 98.6% accuracy of the 2014 electoral roll). The second sample frame consisted of 1,500 people who were listed on the Electoral Roll as being of Maori ancestry, aged between 18-60 years of age, and currently residing in New Zealand. A total of 92 participants responded to this booster sample (response rate = 6.2% adjusting for electoral roll accuracy).

Sampling Procedure -- NZAVS Time 6 (2014)

The Time 6 (2014) NZAVS contained responses from 15,820 participants (15,740 retained from one or more previous wave, and 82 unmatched participants or unsolicited opt-ins). The sample retained 3,728 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 57.2% over five years). The sample retained 14,878 participants from the full Time 5 (2013) sample (a retention rate of 81.5% from the previous year). Participants who provided an email address were first emailed and invited to complete an online version if they preferred. Participants who did not complete the online version (or did not provide an email) were then posted a copy of the questionnaire, with a second postal follow-up two months later. We staggered the time of contact, so that participants who had completed the previous wave were contacted approximately one year after they last completed the questionnaire. As described in the Time 5 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from

the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw. We also emailed participants an online pamphlet containing a series of video interviews with the researchers summarizing different research findings.

Sampling Procedure -- NZAVS Time 7 (2015)

The Time 7 (2015) NZAVS contained responses from 13,942 participants (13,941 retained from one or more previous wave, and 1 unmatched participant or unsolicited opt-ins). The sample retained 3,344 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 51.3% over five years). The sample retained 12,550 participants from the full Time 6 (2014) sample (a retention rate of 79.3% from the previous year). Participants who provided an email address were first emailed and invited to complete an online version if they preferred. Participants who did not complete the online version (or did not provide an email) were then posted a copy of the questionnaire, with a second postal follow-up two months later. We staggered the time of contact, so that participants who had completed the previous wave were contacted approximately one year after they last completed the questionnaire. As described in the Time 5 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw.

Sampling Procedure -- NZAVS Time 8 (2016)

The Time 8 (2016) NZAVS contained responses from 21,936 participants (13,781 retained from one or more previous wave, 7,667 new additions from booster sampling, and 488 unmatched participants or unsolicited opt-ins). The sample retained 3,347 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 51.4%). The sample retained 11,933 participants from the full Time 7 (2015) sample (a retention rate of 85.6% from the previous year). Participants who provided an email address were first emailed and invited to complete an online version if they preferred. Participants who did not complete the online version (or did not provide an email) were then posted a copy of the questionnaire, with a second postal follow-up two months later. We staggered the time of contact, so that participants who had completed the previous wave were contacted approximately one year after they last completed the questionnaire. As described in the Time 4 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw.

To boost sample size and increase sample diversity for subsequent waves, a booster sample was also conducted by selecting people from the New Zealand electoral roll. As with previous booster samples, sampling was conducted without replacement (i.e., all

people included in previous sample frames were identified and removed from the 2016 roll). The sample frame consisted of 80,000 people aged from 18-65 randomly selected from the 2016 New Zealand Electoral Roll, who were currently residing in New Zealand (one can be registered to vote in New Zealand but living overseas). The New Zealand Electoral Roll contains participants' date of birth (within a one-year window), and we limited our frame to people who 65 or younger, due to our aim of retaining participants longitudinally. A total of 7667 participants responded to this booster sample (response rate = 9.7% when adjusting for the 98.6% accuracy of the 2016 electoral roll).

Sampling Procedure -- NZAVS Time 9 (2017)

The Time 9 (2017) NZAVS contained responses from 17,072 participants (16,931 retained from one or more previous wave, and 141 unmatched participants or unsolicited opt-ins). The sample retained 2,771 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 42.5% over five years). The sample retained 15,784 participants from the full Time 8 (2016) sample (a retention rate of 72.0% from the previous year). Participants who provided an email address were first emailed and invited to complete an online version if they preferred. Participants who did not complete the online version (or did not provide an email) were then posted a copy of the questionnaire, with a second postal follow-up two months later. We staggered the time of contact, so that participants who had completed the previous wave were contacted approximately one year after they last completed the questionnaire. We offered a prize draw for participation (five draws each for \$1000 grocery vouchers, \$5000 total prize pool). All participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw.

The retention rate from Time 8 to Time 9 of 72% was notably lower than the ~80% achieved in previous few years. We had opted not to phone non-respondents during the Time 9 wave of data collection, and instead decided to let one-year past and then intensify phoning of non-respondents the following year. We decided on this approach in the hope that it might reduce 'contact fatigue', and hence increase the recovery rate and give more time (two years) for intermittent and non-respondents to become re-enthused about participating in the study. We thus decided to risk a lower retention rate in Time 9 with the hopes of getting a bounce back in the recovery rate during the following wave of data collection amongst these most hard-to-retain participants (see Satherley et al. 2015, for a detailed analysis of the demographic and individual difference factors predicting retention, non-response and intermittent response). Our decision to reduce the risk of contact fatigue during Time 9 was also compounded by a major telecommunications carrier in New Zealand opting to discontinue their email servers in November 2017. Many of our participants had used these email services, and hence a large number of the emails in our database were rendered invalid. This increased the subsequent load on attempting to contact participants via postal mail and phone until contact could be made and our email database could be updated with participants' new email addresses.