The New Zealand Journal of Psychology

Volume 50, Issue 2 (August), 2021
ISSN:1179-7924

New Zealand Psychological Society
Rōpū Mātai Hinengaro o Aotearoa

New Zealand Journal of Psychology

EDITOR

Marc Wilson

School of Psychology
Victoria University of Wellington

ASSOCIATE EDITORS

Neville Blampied

Department of Psychology
University of Canterbury

Janet Leathem

School of Psychology
Massey University at Wellington

Harlene Hayne

Department of Psychology
University of Otago

Michael O'Driscoll

Department of Psychology
University of Waikato

BOOK REVIEW EDITOR

Iris Fontanilla

Auckland DHB

© The material published in this issue is copyright to the New Zealand Psychological Society.

Publication does not necessarily reflect the views of the Society

The New Zealand Journal of Psychology *is published online in three editions a year (articles will be posted as they become available) by the* **New Zealand Psychological Society Inc.**

Content may include manuscripts and shorter research notes in any substantive area of Psychology, and book reviews. Submitted manuscripts may be (1) empirical, (2) reviews of the literature, or (3) discussions of theoretical/conceptual frameworks of relevance to the practice of Psychology. Manuscripts will be considered for publication if they (a) include data collected from New Zealand samples, or (b) discuss the relevance of issues contained in the manuscript to the New Zealand social and cultural context, or to the practice of Psychology in this country. For further clarification of these requirements, please contact the Editor.

See **Instructions to Authors** *at:* https://www.psychology.org.nz/publications-media/new-zealand-journal-of-psychology/#author-guidelines *for preparation and submission of manuscripts and material.*

Subscriptions

Subscription to the Journal is included in Membership dues for the Society. Non-members may subscribe at NZ$55.00 per volume/year (three issues), plus GST (NZ only). Subscription and advertising enquiries should be addressed to the:

Business Manager
New Zealand Journal of Psychology
NZ Psychological Society Inc.
(contact details below)

Production, Printing & Distribution
Is managed by National Office

NZ Psychological Society Inc.
P.O. Box 10536, The Terrace,
Wellington 6143,
New Zealand
Ph (04) 473 4884
Email: office@psychology.org.nz

Editor's Introduction

Firstly, Volume 50, Issue 2, presents six papers, including a pair of studies by Aoroaro Tamati and colleagues presenting, first, the development of culturally appropriate instruments that are then used in a 10-month study of the development of positive behaviours among Māori preschoolers. I am particularly appreciative of the effort of those who reviewed this manuscript, given that it required commitment to evaluating each separately and as a pair. Continuing this developmental theme, this issue also includes Azeem and colleagues' evaluation of the impact of reading age-appropriate books describing inter-race friendships. Continuing this 'evaluation' theme, Hayward and colleagues present a qualitative analysis of focus groups of young women following a sexual violence prevention workshop. Additionally, this issue includes a NZAVS-based longitudinal investigation of the stability of political preference based on more than 5,000 New Zealanders, and a test of a key theoretical claim derived from Thomas Joiner's Interpersonal Theory of Suicide by Mason and colleagues. This last is at least as important for results relating to that key theoretical claim, as it is for the finding that around a quarter of the convenience sample of university students meet the recommended criterion for clinical concern around suicidal ideation. What these all have in common is not that the University of Otago has close to colonised this issue of the Journal, but the significant amount of work that all of these projects represent.

Secondly, we note the passing of two longstanding members of Aotearoa's psychology community - Emeritus Professor Jim Flynn and Emeritus Professor Tony Taylor. If readers have a glance through the archives of the NZJP available on the Society website, they'll see that Tony Taylor has contributed numerous times to these pages, and since the very first volume half a century ago.

Finally, we renew our call for papers to be considered as part of our upcoming special issue on psychological perspectives on environment, climate and sustainability.

CALL FOR PAPERS: Environment, climate and sustainability

We invite submissions for an upcoming special issue of the *New Zealand Journal of Psychology* focused on environment, climate, and sustainability.

We welcome contributions in the form of scholarly reviews, empirical research whether qualitative or quantitative in nature, commentaries relating to practice relevant to psychologists working Aotearoa New Zealand, and other relevant contributions. Contributions should be broadly relevant to the themes of environment, climate and sustainability, but more specific examples might include:

• Psychological and behavioural aspects of people and nature
• Ecological consequences of human actions
• Perception of, and behavioural responses to, environmental risks and hazards
• Māori and indigenous perspectives on the environment and it's impacts
• Effects of environment (and perception of the environment) on human cognition and health
• Theories of environment-relevant behaviour, values, norms, attitudes, and personality
• Psychology of sustainability and climate change
• Psychological practice in the context of environment, climate and sustainability

Special issue Editors are drawn from the New Zealand Psychological Society's Climate Psychology Task Force and include Brian Dixon, Jackie Feather, Natasha Tassell-Matamua, and Marc Wilson. For further information about the Society's Climate Change initiatives please visit the Society website.

The deadline for submissions is ongoing, but likely to close by November 1st, 2021.

Consistent with the imperative of the Journal, *any* submission must clearly articulate relevance in the context of Aotearoa New Zealand. Information about the Journal, and general author guidelines can be found here.

Marc Wilson

New Zealand Journal of Psychology

Volume 50, Number 2 (August), 2021

CONTENTS

Title	Authors	Page
A Test of the Three-Way Interaction of the Interpersonal Theory of Suicide in an Aotearoa New Zealand University Sample	*Andre Mason, Benjamin C. Riordan, Kyungho Jang, Sunny C Collings, and Damian Scarf*	4
Stability and change in New Zealanders' political party support	*Nicole Satherley, Danny Osborne, and Chris G. Sibley*	9
He Piki Raukura: Assessing Ao Māori developmental constructs – Part I: Reliability of novel strengths-based measures among preschool Māori children	*Aroaro Tamati, Gareth J. Treharne, Reremoana Theodore, Mihi Ratima, Erana Hond-Flavell, Will Edwards, Ruakere Hond, and Richie Poulton*	22
He Piki Raukura: Assessing Ao Māori developmental constructs – Part II: Mapping positive change over 10 months among preschool Māori children	*Aroaro Tamati, Gareth J. Treharne, Jesse Kokaua, Reremoana Theodore, Mihi Ratima, Erana Hond-Flavell, Will Edwards, Ruakere Hond, Richie Poulton*	35
Residual Rape Myth Acceptance among Young Women Who Have Recently Completed a Sexual Violence Prevention Workshop	*Madeline Hayward, Gareth J. Treharne, Nicola Liebergreen, Katie Graham and Melanie Beres*	46
Measuring Racial Prejudice in New Zealand Pre-schoolers and testing an intervention to reduce the same using brief cross-race friendship picture books	*Anita A. Azeem, John A. Hunter and Ted Ruffman*	55

© This material is copyright to the New Zealand Psychological Society. Publication does not necessarily reflect the views of the Society.

A Test of the Three-Way Interaction of the Interpersonal Theory of Suicide in an Aotearoa New Zealand University Sample

Andre Mason[1], Benjamin C. Riordan[2], Kyungho Jang[1], Sunny C Collings[3], and Damian Scarf[1]

[1]Department of Psychology, University of Otago, Otago, New Zealand.
[2]Centre for Alcohol Policy Research, School of Psychology and Public Health, La Trobe University, Melbourne, Australia
[3]Faculty of Health, Victoria University, Wellington, New Zealand

A central hypothesis of the Interpersonal Psychological Theory of Suicide is that thwarted belonging, perceived burdensomeness, and fearlessness about death interact to predict greater risk of suicide. The current study sought to address the limited research addressing this interaction within university students. Undergraduate students (n = 377) were invited to complete an online survey containing questions related to the study's key concepts. No evidence was found to support the three-way interaction between thwarted belonging, perceived burdensomeness, and acquired capability for suicide, on suicidal behaviour. Further research is required to validate the Interpersonal Psychological Theory of Suicide among university students.

Keywords: *Suicide; Interpersonal Psychological Theory of Suicide; Perceived Burdensomeness; Thwarted Belonging; Fearlessness about Death*

Introduction

At its inception, Joiner's Interpersonal Psychological Theory of Suicide (Joiner, 2005; Van Orden et al., 2010) was a unique approach to understanding suicide. In contrast to classical theories, Joiner proposed that both suicidal thoughts, *and* the capability to act on those thoughts, are necessary for a potentially lethal suicide attempt. As such, the factors that lead to the development of suicidal ideation were viewed as being distinct from those that facilitate the progression from ideation to attempt. This approach, now known as the ideation-to-action framework (Klonsky & May, 2014), serves as the foundation for contemporary models of suicide.

Broadly, the Interpersonal Psychological Theory of Suicide outlines three key components that contribute to an individual's suicidal thoughts and behaviours; thwarted belonging, perceived burdensomeness, and acquired capability for suicide. The combination of thwarted belonging and perceived burdensomeness is thought to lead to the emergence of suicidal ideation. On their own, however, these feelings are unlikely to result in a lethal suicide attempt. Instead, Joiner argues that the progression to a potentially lethal suicide attempt is facilitated by an acquired capability for suicide; one aspect of which is an individual's capacity to overcome the natural fear of death. Thus, subjective levels of low belonging, high burdensomeness, and high fearlessness about death are hypothesised as providing the conditions required for a suicide attempt.

Despite evidence recognising university students as a vulnerable population for suicide throughout the world (Mortier et al., 2018), only two studies, both using American undergraduates, have demonstrated the three-way interaction in a university student population (Becker et al., 2020; Davidson et al., 2010) . The current study seeks to replicate these earlier findings within Aotearoa New Zealand.

Aotearoa New Zealand's youth suicide rates are consistently identified as being the highest in the world (OECD, 2021). Annual provisional suicide statistics indicate that between July 2019 and June 2020, 119 youth aged between 15-24 died by suicide with the rate per 100,000 being greater in males, relative to females. Furthermore, the rates of suicide for Māori were 1.68, 2.86, and 2.56 times higher than those identifying as 'European and other', Pacific peoples, and Asian (Ministry of Justice, 2020). Given the clear impact of ethnicity, and therefore, the likely impact of culture, it is important to determine whether the proposed interaction holds in other countries that may have a more culturally diverse population.

Consideration regarding the influence of culture within the context of the Interpersonal Psychological Theory of Suicide is particularly important given that different cultures may have greater emphasis on core components within the model. For example, te ao Māori typically adopts a more holistic approach that acknowledges the inter-relatedness of all elements infused by *mauri* (life force). As such, the salience of concepts like thwarted belonging or perceived burdensomeness may be heightened for Māori, and, therefore, contribute to an elevated risk of suicide. The current study proposes to test the three-way interaction between thwarted belonging, perceived burdensomeness, and fearlessness about death in a general Aotearoa New Zealand university sample. In alignment with Becker et al. (2020), we hypothesise that the three-way interaction would predict significantly higher scores on a composite measure of suicide risk.

METHODS

Participants

Participants were 377 undergraduate psychology students enrolled at a New Zealand university. Participants ranged in age from 18–25 years (*M*=19.59, *SD* = 1.18) and predominantly identified as female (*n* = 316, 83.8%; 60 males, 15.9%; 1 non-binary, 0.8%). The majority of the participants self-identified as New Zealand European (75.6%; *n* = 285); the remaining participants identified as New Zealand Māori (*n* = 20), Samoan (*n* = 4), Cook Island Māori (*n* = 1), Tongan (*n* = 2), Chinese (*n* = 10), and Indian (*n* = 8). Forty-seven (12.5%) stated specific ethnicities which generally fell into Asian or European regions.

Measures

Thwarted Belonging & Perceived Burdensomeness: The 15-item Interpersonal Needs Questionnaire (INQ; Van Orden et al., 2012) measured participants' sense of belonging and perceived burdensomeness. Participants were asked to respond to each item using a seven-point Likert scale (not at all true for me = 1 to very true for me = 7) with higher scores indicating higher levels of thwarted belonging and perceived burdensomeness (Cronbach's αs = .88 (*se* = .01; 95% *CI*: .85-.90) and .93 (*se* = .01, 95% *CI*: .90 - .94) for thwarted belonging and perceived burdensomeness, respectively).

Fearlessness About Death: The Acquired Capability for Suicide Scale – Fearlessness About Death (ACSS-FAD; Ribeiro et al., 2014) contains seven-items rated on a five-point scale (not at all like me = 0 to very much like me = 4). Higher scores represent greater levels of fearlessness (Cronbach's α = .80, *se* = .02, 95% *CI*: .77-.83).

Suicide Risk: The Suicidal Behaviour Questionnaire-Revised (SBQ-R; Osman et al., 2001) assesses four dimensions of suicidality (e.g. lifetime ideation/attempt) and responses can be used to identify at-risk participants with high sensitivity (93%) and specificity (95%; Osman et al., 2001). Higher scores indicate greater levels of suicide risk (Cronbach's α = .83, *se* = .02, 95% *CI*: .79 - .86). SBQ-R is the most commonly used composite measure of suicidality (Chu et al., 2017).

Depression: The seven-item depression subscale of the Depression Anxiety Stress Scales-21 (Lovibond & Lovibond, 1995) assessed depressive symptoms, rated on a four-point likert scale (never = 0 to almost always = 3). Higher scores indicate greater depressive symptoms (Cronbach's α = .89, *se* = .01, 95% *CI*: .87 - .91).

Procedure

This study was approved by the University of Otago Ethics Committee (H20/029) and the hypotheses were pre-registered with the Open Science Framework (https://osf.io/cj2mr/). Undergraduate psychology students were invited to participate for course credit. Students who expressed interest in taking part were emailed about the study and, if they chose to participate, were asked to complete a brief online survey through Qualtrics. Electronic consent was obtained, and every student was given the contact details for the research team, a student health facility, and emergency psychiatric services, in addition to a list of other support services and telephone helplines.

ID numbers were obtained for the purposes of being able to identify students who may be a risk to themselves. If a student met this threshold, the study PI searched the student's ID number within the university database to associate with the name and email address of the student (this could only be done by the PI). It is important to emphasize that students were only identified if they indicated high levels of possible risk and this was clearly articulated to the students within the information sheets.

Table 1. Means, standard deviations, and correlations with confidence intervals

	Age	Gender	Ethnicity	Depressive Symptoms	TB	PB	FAD	SBQ-R
Gender	.04 (-.06, .14)							
Ethnicity	.20** (.10, .29)	.06 (-.04, .16)						
Depressive Symptoms	-.01 (-.11, .09)	-.02 (-.12, .08)	.07 (-.03, .17)					
TB	.02 (-.09, .12)	-.03 (-.13, .07)	.10* (.00, .20)	.68** (.62, .73)				
PB	-.04 (-.14, .06)	.04 (-.06, .14)	.05 (-.05, .15)	.64** (.58, .70)	.60** (.53, .66)			
FAD	.06 (-.04, .16)	-.13* (-.22, -.03)	-.06 (-.16, .04)	.03 (-.07, .13)	.02 (-.08, .12)	.04 (-.07, .14)		
SBQ	.05 (-.05, .15)	-.05 (-.15, .05)	.05 (-.05, .15)	.52** (.44, .59)	.47** (.38, .54)	.56** (.49, .63)	.13* (.03, .23)	
Mean	19.59			9.11	24.45	10.53	13.12	7.31
SD	1.18			8.47	9.18	5.74	6.14	3.54
Range	18 - 25			0 - 40	9 - 61	6 - 42	0 – 28	4-21

PB = Perceived Burdensomeness, TB = Thwarted Belonging, FAD = Fearlessness About Death, SBQ-R = Suicide Behaviour Questionnaire – Revised

**p < .05, ***p < .001

Table 2. Multiple regression testing the three-way interaction between perceived burdensomeness, thwarted belonging, and fearlessness about death, on suicidal risk, including depressive symptoms as a covariate.

DV = SBQ-R	*b*	SE	*B (95% CI)*	*t*	*p*
			$F(8, 368) = 28.35$, $p < .001$, $Adj\ R^2 = .37$ (95% CI: .29, .44)		
Depressive symptoms	0.10 (0.05, 0.15)	.03	0.23 (0.11, 0.35)	3.72	<.001
PB	0.14 (-0.27, 0.54)	.20	0.22 (-0.43, 0.87)	.66	.51
TB	0.02 (-0.14, 0.17)	.08	0.04 (-0.37, 0.45)	.19	.85
FAD	-0.08 (-0.34, 0.17)	.13	-0.14 (-0.59, 0.30)	-.64	.52
PB X TB	0.00 (-0.01, 0.01)	.01	0.04 (-0.84, 0.93)	.10	.92
PB X FAD	0.01 (-0.01, 0.04)	.01	0.39 (-0.43, 1.20)	.94	.35
TB X FAD	0.00 (-0.01, 0.01)	.01	0.21 (-0.38, 0.79)	.70	.49
PB X TB X FAD	-0.00 (-0.00, 0.00)	.00	-0.31 (-1.27, 0.64)	-.64	.52

PB = Perceived Burdensomeness, TB = Thwarted Belonging, FAD = Fearlessness About Death, SBQ-R = Suicide Behaviour Questionnaire – Revised

Power Analysis and Analytic Strategy

A simulation based on previous findings by Becker et al. (2020) was used to estimate sample size required to detect the three-way interaction between perceived burdensomeness, thwarted belonging, and fearlessness about death (estimated effect size = .07) on suicide risk. The target was at least 80% power. All analyses were conducted using R (version 4.0.0; R Core Team, 2020). The simulation revealed that a minimum sample size of n = 143 would yield >80% power while a minimum sample size of n = 333 would yield >99% power.

Initially, descriptive data and Pearson's correlations were calculated using the psych (Revelle, 2018), MBESS (Kelley, 2020) and Hmisc (Harrell Jr, 2020) packages in R (version 4.0.0; R Core Team, 2020). Multiple linear regression using R's lm.beta package (Behrendt, 2014) was then used to test the three-way interaction between perceived burdensomeness, thwarted belonging, and fearlessness about death, on suicidal risk.

RESULTS

Preliminary Data Analyses

Descriptive data and Spearman's correlations were calculated using the psych (Revelle, 2018) and Hmisc (Harrell Jr, 2020) packages in R (version 4.0.0; R Core Team, 2020). With the exception of age (kurtosis = 4.93) and perceived burdensomeness (4.26), all data were within acceptable limits of skew (<2.0) and kurtosis (< 4.0). Based on the cut-off scores for the SBQ-R total score, as outlined by Osman et al (2001), 100 (26.5%) of students met the threshold for 'at-risk' of suicide; this proportion is somewhat comparable to Robinson et al.,

Table 3. Multiple regression testing the three-way interaction between perceived burdensomeness, thwarted belonging, and fearlessness about death, on suicide risk

DV = SBQ-R	*B (95% CI)*	SE	*B (95% CI)*	*t*	*p*
			$F(7, 369) = 29.4$, $p < .001$, $Adj\ R^2 = .39$ (95% CI: .27, .42)		
PB	.16 (-0.24, 0.57)	.21	0.26 (-0.39, 0.92)	.79	.43
TB	.05 (-0.11, 0.21)	.08	0.13 (-0.28, 0.55)	.62	.53
FAD	-.07 (-0.33, 0.19)	.13	-0.13 (-0.58, 0.32)	-.55	.58
PB X TB	.00 (-0.01, 0.01)	.01	0.14 (-0.76, 1.05)	.31	.75
PB X FAD	.01 (-0.01, 0.04)	.01	0.41 (-0.42, 1.23)	.97	.34
TB X FAD	.00 (-0.01, 0.01)	.01	0.21 (-0.38, 0.80)	.69	.49
PB X TB X FAD	-.00 (-0.00, 0.00)	.00	-0.39 (-1.36, 0.58)	-.79	.43

PB = Perceived Burdensomeness, TB = Thwarted Belonging, FAD = Fearlessness About Death, SBQ-R = Suicide Behaviour Questionnaire – Revised

(2021) who found that in a large New Zealand adolescent sample, 19.3% met the clinical cut-off. Furthermore, the observed mean (7.32, *SD* = 3.54) is comparable to that observed by Becker et al. (2020) in an American university sample (*M* = 5.17; *SD* = 2.82).

Additionally, the mean scores recorded on the depressive subscale of the DASS-21 (*M* = 9.11, *SD* = 8.47) correspond to the upper bounds of the "normal" range; 151 (25%) of students met the threshold for moderate depression or higher (Lovibond & Lovibond, 1995). Neither age, gender, nor ethnicity were significantly correlated with suicide risk and were not considered further. As shown in Table 1, moderate to strong correlations were found among depression, thwarted Robinson et al. (2021) belonging, perceived burdensomeness, and suicidal risk. Additionally, fearlessness about death was weakly correlated with suicide risk.

Primary Data Analyses

Multiple linear regression using R's lm.beta package (Behrendt, 2014) was used to test the three-way interaction between perceived burdensomeness, thwarted belonging, and fearlessness about death, on suicidal risk. Given its significant relationship with suicide risk, depressive symptoms were also included in this model. Only depressive symptoms significantly predict suicidal risk; there was no evidence for the presence of a three-way interaction (see Table 2). Furthermore, no effects emerged when depression was removed from the model (see Table 3).

DISCUSSION

In the current study, we found no evidence for a three-way interaction between thwarted belonging, perceived burdensomeness, and fearlessness about death on suicidal behaviour in Aotearoa New Zealand undergraduates. Although the interaction is a core tenet of the Interpersonal Psychological Theory of Suicide (Joiner, 2005; Van Orden et al., 2010), there is limited evidence regarding this interaction within university student populations outside of the US. Given evidence that cites country-specific contextual factors as a possible limitation for global multi-level interventions for suicide (Collings et al., 2018) and recognises students worldwide as a vulnerable population (Mortier et al., 2018), our failure to replicate earlier results suggests more work is needed in this area. It is imperative, therefore, that research continues to empirically evaluate theoretical models of suicide in different populations and cultural contexts, especially given the over-representation of Māori within New Zealand's suicide statistics

As noted above, our findings are consistent with a growing body of literature demonstrating that the interactions within the Interpersonal Psychological Theory of Suicide are somewhat tenuous (Chu et al., 2017; Ma et al., 2016). For example, in their systematic review, Ma et al. (2016) reported that only three of seven studies displayed a significant three-way interaction between thwarted belonging, perceived burdensomeness, and fearlessness about death. Moreover, only one of these studies examined university students.

The primary limitation of the current study is its cross-sectional nature, which limits possibilities for causal or temporal inferences. To address this, future studies could employ a longitudinal design similar to Forkmann et al. (2020). Additionally, while the current study is, to our knowledge, the first to explore the interactions within the Interpersonal Psychological Theory of Suicide in an Aotearoa New Zealand university context, it is important to note that our findings relate to undergraduate psychology students at a single New Zealand university, most of whom were female. Although psychology students are drawn from a number of different majors and degrees, and is one of the most popular courses at the host university, the homogeneity of the present sample limits our ability to comment on possible risk and protective factors for suicidal thoughts and behaviours, particularly for those who are most at risk (i.e. males, particularly male Māori youth). Further research that specifically recruits students from a more diverse range of genders, cultures, and academic study as well as from different universities within New Zealand is required, therefore, to advance our understanding in this area. Such scholarship would increase our understanding of factors that may lead to the emergence of suicidal thoughts and behaviours, which, in turn, may help to improve intervention and prevention strategies for suicide in Aotearoa New Zealand.

References

Becker, S. P., Foster, J. A., & Luebbe, A. M. (2020). A test of the interpersonal theory of suicide in college students. *Journal of Affective Disorder*, *260*, 73-76. https://doi.org/10.1016/j.jad.2019.09.005

Behrendt, S. (2014). *lm.beta: Add Standardized Regression Coefficients to lm-Objects* In (Version R package version 1.5-1)

Chu, C., Buchman-Schmitt, J. M., Stanley, I. H., Hom, M. A., Tucker, R. P., Hagan, C. R., Rogers, M. L., Podlogar, M. C., Chiurliza, B., Ringer, F. B., Michaels, M. S., Patros, C. H. G., & Joiner, T. E. (2017). The interpersonal theory of suicide: A systematic review and meta-analysis of a decade of cross-national research. *Psychological Bulletin*, *143*(12), 1313-1345. https://doi.org/10.1037/bul0000123

Collings, S., Jenkin, G., Stanley, J., McKenzie, S., & Hatcher, S. (2018). Preventing suicidal behaviours with a multilevel intervention: a cluster randomised controlled trial. *BMC Public Health*, *18*(1), 140. https://doi.org/10.1186/s12889-018-5032-6

Davidson, C. L., Wingate, L. R., Slish, M. L., & Rasmussen, K. A. (2010). The great black hope: Hope and its relation to suicide risk among African Americans [Article]. *Suicide and Life-Threatening Behavior*, *40*(2), 170-180. https://doi.org/10.1521/suli.2010.40.2.170

Forkmann, T., Glaesmer, H., Paashaus, L., Rath, D., Schonfelder, A., Stengler, K., Juckel, G., Assion, H. J., & Teismann, T. (2020). Interpersonal theory of suicide: prospective examination. *British Journal of Psychiatry Open*, *6*(5). https://doi.org/10.1192/bjo.2020.93

Harrell Jr, F. E. (2020). *Hmsic: Harrell Miscellaneous*. In (Version R package version 4.4-0)

Joiner, T. E. (2005). *Why people die by suicide*. Harvard University Press.

Klonsky, E. D., & May, A. M. (2014). Differentiating suicide attempters from suicide ideators: a critical frontier for suicidology research. *Suicide and Life Threatening Behavior*, *44*(1), 1-5. https://doi.org/10.1111/sltb.12068

Lovibond, P. F., & Lovibond, S. H. (1995). *Manual for the Depression Anxiety & Stress Scales.* (2 ed.). Psychology Foundation.

Ma, J., Batterham, P. J., Calear, A. L., & Han, J. (2016). A systematic review of the predictions of the Interpersonal-Psychological Theory of Suicidal Behavior. *Clinical Psychology Review*, *46*, 34-45. https://doi.org/10.1016/j.cpr.2016.04.008

Ministry of Justice. (2020). *Annual provisional suicide statistics for deaths reported to the Coroner between 1 July 2007 and 30 June 2020*. https://coronialservices.justice.govt.nz/assets/Documents/Publications/2020-Annual-Provisional-Suicide-Statistics.pdf

Mortier, P., Auerbach, R. P., Alonso, J., Bantjes, J., Benjet, C., Cuijpers, P., Ebert, D. D., Green, J. G., Hasking, P., Nock, M. K., O'Neill, S., Pinder-Amaker, S., Sampson, N. A., Vilagut, G., Zaslavsky, A. M., Bruffaerts, R., Kessler, R. C., & Collaborators, W. W.-I. (2018). Suicidal Thoughts and Behaviors Among First-Year College Students: Results From the WMH-ICS Project. *Joutrnal of the American Academy of Child and Adolescent Psychiatry*, *57*(4), 263-273 e261. https://doi.org/10.1016/j.jaac.2018.01.018

OECD. (2021). *Suicide rates (indicator)* https://doi.org/10.1787/a82f3459-en

Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviours Questionnaire - Revised (SBQ-R): Validation with Clinical and nonclinical samples. *Assessment*, *8*, 443-454. https://doi.org/https://doi.org/10.1177/107319111010080 0409

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. In R Foundation for Statistical Computing.

Revelle, W. (2018). *psych: Procedures for Psychological, Psychometric, and Personality Research* In (Version R package version 1.8.12)

Ribeiro, J. D., Witte, T. K., Van Orden, K. A., Selby, E. A., Gordon, K. H., Bender, T. W., & Joiner, T. E., Jr. (2014). Fearlessness about death: the psychometric properties and construct validity of the revision to the acquired capability for suicide scale. *Psychological Assessment*, *26*(1), 115-126. https://doi.org/10.1037/a0034858

Robinson, K., Garisch, J. A., & Wilson, M. S. (2021). Nonsuicidal self-injury thoughts and behavioural characteristics: Associations with suicidal thoughts and behaviours among community adolescents. *Journal of Affective Disorders*, *282*, 1247-1254. https://doi.org/10.1016/j.jad.2020.12.201

Van Orden, K. A., Cukrowicz, K. C., Witte, T. K., & Joiner, T. E. (2012). Thwarted belongingness and perceived burdensomeness: construct validity and psychometric properties of the Interpersonal Needs Questionnaire. *Psychological Assessment*, *24*(1), 197-215. https://doi.org/10.1037/a0025358

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E., Jr. (2010). The interpersonal theory of suicide. *Psychological Review*, *117*(2), 575-600. https://doi.org/10.1037/a0018697

Corresponding Author
Andre Mason
Department of Psychology, PO Box 56, University of Otago, Dunedin 9054, New Zealand.
Email: andre.mason@postgrad.otago.ac.nz

Stability and change in New Zealanders' political party support

Nicole Satherley, Danny Osborne, and Chris G. Sibley
School of Psychology, University of Auckland

Like in many multiparty contexts, questions remain about the nature of stability and change in New Zealanders' partisanship. Here, we add to the literature by systematically analysing the stability and change in party support and voting using a large longitudinal national probability sample of New Zealand adults (Ns = 5,449–9,845). Support (measured from strongly oppose to strongly support) for two major and two minor parties was generally stable over two-year (.58-.84) and five-year (.51-.77) test-retest periods. Political identity centrality moderated this association such that party support was more stable among those whose political beliefs were central to their sense of self. Markov models identified relatively stable voting patterns for major parties (National and Labour), with low probabilities of switching between these parties at elections in 2014 (.04-.07) and 2017 (.04 -.12). These results suggest support for political parties (measured via behaviour and attitudes) is generally stable in New Zealand.

Keywords: *Partisanship, Political Identity, Party Support, Voting Stability, Multiparty Systems*

Introduction

Political parties have long received attention for their role in shaping voters' political attitudes and voting behaviour. In their seminal work on the topic, Campbell, Converse, Miller, and Stokes (1960) documented that political partisanship was an enduring and rigid psychological attachment to a political party that shapes political behaviour. Such a conception has led many to characterise party identification as an "unmoved mover" (e.g., see Green & Palmquist, 1990)—and for good reason. Party identification in the United States (U.S.) is highly stable over time (Green & Palmquist, 1994; Green, Palmquist, & Schickler, 2002). This feature of partisanship helps to explain its capacity to shape both policy preferences (Carsey & Layman, 2006) and political values (Goren, Federico, & Kittilson, 2009).

Despite the important outcomes of stable political partisanship, there has been only limited focus on its stability in New Zealand. Indeed, questions over the stability of party attitudes within *multi*party systems have persisted over the years (e.g., see Green et al. 2002; Johnston 2006; Thomassen & Rosema, 2009). Whereas identification with one party in two-party systems like the U.S. is generally associated with low or no identification with the other party (Schickler & Green, 1997), citizens of multiparty systems may identify with more than one party (Green et al., 2002; Schickler & Green, 1997). Multiple parties may also adopt similar ideological stances (Johnston 2006), thereby obscuring differences between political platforms. For example, support for one party may be readily relinquished in favour of another when ideological disagreements arise in multiparty systems because there are viable, ideologically similar, substitutes. Alternatively, citizens may express *greater* stability in their party preferences if more nuanced ideological differences between parties make it easier for voters to find the 'best fitting' party.

In this study, we aim to provide insights into the stability of New Zealanders' political party support across the adult lifespan. Using a large longitudinal national probability sample, we examine the stability of party support (i.e., an attitudinal indicator of partisanship) across two- and five-year periods, and also investigate the hypothesis that political identity centrality (i.e., how important one's political beliefs are to their sense of self) moderates these stability estimates. To complement these analyses, we then examine stability and change in party votes (i.e., behavioural indicators of partisanship) across three successive national elections (2011, 2014, and 2017) over the same period. In doing so, our analyses increase understanding of the stability and centrality of party attitudes of New Zealanders, as well as voters in other multiparty systems.

Partisan stability in New Zealand and other multiparty systems

Partisanship in the United States is often assessed via party identification—that is, how strongly one identifies with a given political party (see Campbell et al., 1960). Yet limitations and issues with using party identification as a measure of partisanship have long been noted (e.g. Thomassen & Rosema, 2009). For example, difficulties in translating the concept of identity, and adapting the question wording for different parties has resulted in what Johnston (2006) refers to as a "measurement swamp" in the cross-national literature. The original measure also does not account for identification with multiple parties, which is more common in multiparty systems (see Green et al., 2002; Schickler & Green, 1997; Weisberg, 1980), or 'negative partisanship' based on dislike of parties (e.g., Caruana, McGregor, & Stephenson, 2015; Mayer, 2017; Rose & Misher, 1998).

Research on the stability of party attitudes in multiparty systems has consequently utilized inconsistent measures and produced disparate results. For example, Kuhn (2009) found low levels of stability in voting intentions using Swiss Household Panel data from 1999-2007, as only 18% of respondents reported voting for the same party in each wave (although a further 27% stated

either the same party preference or no preference across waves). Schmitt-Beck, Weick, and Christoph (2006) also assessed party identification over an extended period (1984-2001) in a German panel study. They found that only around 24% of respondents remained committed to the same party in every wave, casting doubt on the long-term stability of partisanship in multiparty systems. However, Schmitt-Beck et al. also noted that most of the instability in party identification occurred due to movement to-and-from the same party and identification as an "independent" voter. In other words, respondents were unlikely to switch between different parties, but rather, varied in the constancy in which they identified with a single party over time.

Using a unique scale assessing social identification with political parties, Huddy, Bankert, and Davies (2018) found evidence of stability in the strength of people's partisan identity in a multiparty system using data from the British Election Study. Their standardized estimates of stability ranged from .24 (over a period of 18 months) to .79 (over a period of 6 months), and were somewhat higher among Conservative (vs. Labour) Party identifiers. Others have found even stronger evidence of the stability of party attitudes in multiparty systems. For example, Richardson (1991) found the stability of party identification in Britain, the Netherlands, and West Germany to range between $r = .66 – .79$ over one to three years. Similarly, Schickler and Green (1997) found that party identification was highly stable in Britain, Germany, and Canada, reaching stability estimates comparable to the U.S. Using dummy-coded variables indicating whether or not a respondent most strongly identified with a given party, they found R^2 values generally greater than .83 over periods ranging from 5 months to 4 years. Thus, some studies suggest that partisanship can indeed be highly stable in multiparty contexts.

In New Zealand, extant research on partisan stability is scarce, and mostly predates the 1996 shift to mixed-member proportional (MMP) representation, which ushered in a viable multiparty political era (see Vowles, 2005). Dalton and Weldon (2007) identified similar rates of feeling close to a party in New Zealand (56.3%) compared to the U.S. (57%) between 1996 – 2000, suggesting similar rates of partisanship. Lamare (1984) also found that party identification nullifies the impact evaluations of policies, political issues, and party candidates have on vote choice. Yet, other studies suggest party identification may not translate to stable partisan behaviours over time. Aimer (1989) adapted the party identification scale for the New Zealand context and found high rates (86%) of identification during the 1987 election. Identification also corresponded closely to vote choice (79% rate of matching), yet only 54% - 66% of Labour and National Party identifiers repeatedly voted for their party across three successive elections. Leithner (1997) also found generally low rates (roughly 60%) of loyal voting for the National and Labour parties from 1935 – 1987. More recently, Vowles (2016) identified a decline in the percentage of New Zealanders reporting strong or close party identification since 2005, which was just below 40% in 2014. Vowles' data also showed a general increase in voter volatility from 1972-2014, with around 40% of NZES participants switching their vote (between parties or to-and-from abstention) in 2014.

Although multiparty systems possess unique qualities that may influence stability (i.e., more parties and potentially greater ideological diversity), these factors can vary across contexts. Indeed, this may help to explain the variability in stability estimates found across studies. Analysing data from the Comparative Study of Electoral Systems, Dassonneville, Blais, and Dejaeghere (2015) showed that the effective number of parties in a system moderates the volatility in party votes. Switching votes between parties occurred more often in systems with a greater number of parties *regardless* of how much participants liked the party for which they had previously voted (although the number of parties made no difference when participants expressed strong dislike for the party for which they previously voted). Given the effective number of parties in New Zealand has been relatively lower and more stable in recent years (i.e., 3.3 – 2.9 in 2014 and 2017 respectively, Vowles, 2018), this may suggest relatively high levels of stability in party support and voting behaviour during the period examined here.

Other factors may also hint at higher levels of stability. For example, partisanship tends to be more prevalent in older systems and is fostered through parental socialization (Dalton & Weldon, 2007). New Zealand's two main political parties (National and Labour) are longstanding in the political landscape (having been established over 85 years ago), which may increase the odds of inter-generational transferal of party preferences. New Zealand's parties are also reasonably well ideologically polarized (Dalton, 2008), which may make it easier for New Zealanders to perceive differences between parties and, therefore, maintain support and voting for a given party. Thus, several factors in multiparty contexts may shape partisanship and the stability of party attitudes and behaviour, and these factors may point toward relatively high stability in New Zealand.

Overview of the present study

We contribute to the literature by examining the stability of political party support across the adult lifespan for four main political parties in New Zealand: The National Party, the Labour Party, the Greens, and New Zealand (NZ) First. The centre-right National Party (in government from 2008 – 2017) and the centre-left Labour Party are the two major parties in New Zealand that traditionally receive the most support during general elections (e.g., the parties received 44.45% and 36.89% of the vote in 2017, respectively). The Greens (6.27% of the 2017 vote), in contrast, are an environmentally focused minor party who also advocate for greater societal equality, and often work closely with the Labour Party. Finally, NZ First (7.20% of the 2017 vote) tends to combine populist sentiment with nationalism (e.g., promoting restrictions to immigration), yet progressively advocate for senior citizens. In this sense, NZ First inconsistently aligns with the left and the right, and has formed governments with both National and Labour (see Vowles, 2018).

Although party identification is most commonly used to assess partisanship, notions of party support are widely used in its description. For example, Bartle and Belluci (2009) define partisanship as "a long-term tendency to

support one party rather than another" (p. 1), whereas Petrocik (2009) asserts that "it represents an expression of support that influences behavior and other party-related attitudes and assessments." (p. 564). As such, we use ratings of support for each party (ranging from strongly support to strongly oppose) as an attitudinal indicator of partisanship. Indeed, measures of party support are useful in that they allow for both negative partisanship (i.e., partisanship defined by opposition toward parties; Caruana et al., 2015) and the presence of simultaneous support for multiple parties. We then assess the stability of party vote as a behavioural indicator of partisanship (see Bartels, 2000, for the strong association between partisanship and voting in the U.S.) and, crucially, whether party support predicts future party vote while adjusting for the effect of prior party vote.

Research has demonstrated the utility of party support measures to examine partisanship in New Zealand. Highlighting the presence of multiple party commitments, Greaves et al. (2015) conducted a latent profile analysis of party support ratings and found that, whereas 14% of the sample exclusively supported the Labour Party, a further 16.4% supported both the Labour Party and the Greens. Moreover, Satherley, Yogeeswaran, Osborne, and Sibley (2018) showed that party support was predictive of party-consistent attitude change at the 2016 flag change referendums, suggesting partisanship is highly central to New Zealanders. Thus, by examining the stability of explicit measures of party support for the first time, our analyses provide further insight into their utility in assessing partisanship in multiparty contexts. Whereas most past research in New Zealand has examined aggregated indices of partisan stability (for example, examining overall net volatility in voting; see, e.g., Leithner, 1997; Vowles, 2016), our analyses also provide a more systematic party-by-party examination of stability including specific patterns of party vote change (i.e., the probabilities of shifting between specific parties, or to-and-from abstention), as well as moderators of these effects, over time.

Although we generally expected people's party support to be stable over time, the amount of stability should vary depending on the party examined. Indeed, past research has found that preferences for minor (versus major) parties are more susceptible to change (Kuhn 2009; Richardson 1991; Schmitt-Beck et al. 2006). Kuhn (2009) argues that the decline in stability for minor parties results from their relatively low power and that support for minor parties tends to fluctuate in relation to particularly salient topics. Greater hostility between ideologically adjacent parties may also promote more out-party negativity among supporters of those parties (Richardson, 1991). Preliminary analyses in New Zealand seem to support this notion. Greaves, Osborne, Sengupta, Milojev, and Sibley (2014) showed that support for the two major parties (namely, the National Party and the Labour Party) was more stable than support for the minor parties (including the Greens), albeit over only a one-year period (i.e., 2009-2010). As such, we expected support for NZ First and the Greens to be less stable than support for the National Party and the Labour Party. In terms of voting behaviour, New Zealand voters may cast strategic votes between their favoured parties based on their anticipated electoral outcome (Bowler, Karp, & Donovan, 2010). This may lead to lower impressions of stability when assessing party votes over time, particularly for concurrent Labour and Green party supporters who may switch their party vote between these parties more often.

We also consider nuances in the stability of party support. The instrumental approach to partisanship (i.e., partisanship as reflecting a 'running tally' of party evaluations; Fiorina, 1981) generally explains stability in terms of enduring expectations about political parties (e.g., Franklin & Jackson, 1983). Yet high stability of partisanship is specifically hypothesised by expressive approaches to partisanship (i.e., partisanship as a deep psychological attachment to a party; Campbell et al., 1960), which argue that the motivated defence of partisan identities foster party-congruent perceptions and attitudes (rather than vice versa; see Bolsen, Druckman, & Cook, 2014). As such, we examined political identity centrality as a moderator of party support and hypothesized that, although party support should generally be stable, those who report higher identity centrality (i.e., those who consider their political beliefs to be more important to their sense of self) should have more stable levels of support. Finally, we also account for age differences in the stability of support in our models, given that the stability of partisanship increases with age (e.g., Green et al., 2002; Sears & Funk, 1999; Stoker & Jennings, 2008). Thus, we generally expected higher levels of stability in older age.

METHODS

Sampling Procedure

We used data from Time 4 (2012), Time 6 (2014) and Time 9 (2017) of the NZAVS, a national probability longitudinal panel study of New Zealand adults sampled from the New Zealand electoral roll. The Time 4 NZAVS consisted of 12,179 participants, 5,107 of whom were obtained through booster sampling and were completing the survey for the first time. The Time 6 NZAVS consisted of 15,820 participants, 9,075 of whom had completed Time 4 (a retention rate of 74.5% of the Time 4 sample). Finally, the Time 9 (2017) NZAVS consisted of 17,072 participants, with 6,776 participants retained from the Time 4 sample (a 3-year retention rate of 55.6%). For each wave, participants received a postal copy of the questionnaire and also had the opportunity to complete the questionnaire online. Participants could also enter a prize draw for grocery vouchers for participation. Non-respondents received multiple reminders to complete the study through phone and email.

Data collection for each wave spanned approximately one year, with Time 4 collection running from 19.09.2012 – 16.09.2013, Time 6 from 21.10.2014 – 19.08.2015, and Time 9 from 13.08.2017 – 17.06.2018. Thus, we refer to each wave according to when data collection began (2012, 2014, and 2017 respectively). Note that the Time 4 (2012) wave assesses participants' party support in 2012/13, as well as their retrospective party vote from the 2011 election (we opted to use the Time 4/2012 wave, rather than the Time 3/2011 wave, as NZ First support was not assessed in the Time 3 questionnaire).

Participants

Participants who completed the 2012 and 2014 waves of the NZAVS had a mean age of 50.71 years (*SD* = 14.56, range = 18 – 93) in 2012, and 63% of the sample were women. Participants were able to report more than one ethnicity, with 89% of the sample reporting New Zealand European, 14% Maori, 4% Asian, and 4% Pacific. Finally, 41% of the sample were religious, 72% were employed, and 80% were born in New Zealand.

Participants who completed the 2012 and 2017 waves of the NZAVS had a mean age of 51.24 years (*SD* = 14.35, range = 18 – 93) in 2012, and 63% of the sample were women. Participants were able to report more than one ethnicity, with 91% of the sample reporting New Zealand European, 12% Maori, 4% Asian, and 3% Pacific. Finally, 40% of the sample were religious, 73% were employed, and 80% were born in New Zealand.

Measures

To assess support for the National Party, Labour Party, Green Party, and NZ First, participants were asked to "Please rate how strongly you oppose or support each of the following political parties", on a scale from 1 (*strongly oppose*) to 7 (*strongly support*). These items were asked in 2012, 2014, and 2017. Voting behaviour was assessed with the item "Did you vote in the New Zealand election in 2014 (2017)" (yes/no) for the Time 6 and Time 9 NZAVS, and "Did you vote in the last (2011) New Zealand general election?" (yes/no) in the Time 4 NZAVS, followed by, "If yes, to which party did you give your party vote?" (open-ended). Political identity centrality was measured with the item, "How important are your political beliefs to how you see yourself?" (1 = *not important*, 7 = *very important*) in both 2014 and 2017.

Model Estimation

To estimate the rank-order stability of party support, we ran a series of models regressing party support in 2014 (for the 2-year estimates) and 2017 (for the 5-year estimates) on the same attitude measured in 2012. These analyses adjusted for the main effect of age (including quadratic and cubic components) and gender measured in 2012, as well as the main effect of identity centrality as measured in either 2014 or 2017 (as this variable was not included in the 2012 wave of the NZAVS). For these analyses, age was scaled by a factor of 10, and age, party support, and political identity centrality were mean-centred. We further included gender by party support and political identity centrality by party support interaction terms, and interaction terms between party support and linear, quadratic, and cubic terms for age. Non-significant higher-order interactions for age were dropped from the model until either a higher-order interaction was significant, or only the linear age interaction term remained (regardless of whether or not it was significant; for similar analytical approaches examining human values and personality traits, see Milfont, Milojev, & Sibley 2016 and Milojev & Sibley, 2014, respectively).[1]

The resulting models allowed us to estimate the simple slope for party support (i.e., the stability coefficient) at each age from 18 – 80 based on the standardized regression terms, as well as at high and low levels of political identity centrality. Although participants' ages ranged to 93 in our analyses, we calculated point estimates for ages up to 80 years because sample sizes were increasingly small at the tail end of the distribution (i.e., less than 1% of the sample was over the age of 80). As such, estimates beyond 80 years of age would have increasingly large standard errors.

RESULTS

Rank-order stability estimates

Table 1 displays correlations between measures across each wave examined. The standardized parameters for the formal regression models assessing the stability of party support are displayed in Tables 2-5. National Party support exhibited the highest stability over both the 2-year and 5-year test-retest periods (.84 and .77, respectively), which reflects the party's status as the most electorally successful party over the period examined. Labour and Green party support displayed comparable levels of stability over the 2-year test-retest period (.71 and .72, respectively), and the 5-year period (.65 and .68, respectively). Although Labour support could be expected to be more stable given its major party status, these findings capture the relatively lower levels of electoral success for the Labour Party over the 5-year period compared to National. Consistent with our hypothesis, NZ First support was the least stable over time. In short, support for political parties tended to be highly stable over time, with some variation depending on major vs. minor party status.[2]

---

[1] We took this approach to adjust for the positive association between age and partisan stability identified in prior research. Because we were not testing for specific hypotheses relating to age (except that stability would generally be higher in older age), we tested for cubic and quadratic terms in an exploratory manner to account for potential non-linear trends. Gender interactions were also included, as the NZAVS tends to over-represent women in particular.

[2] As with any analysis of this nature, it is worth noting that our samples may contain bias that influence our stability estimates (such as, for example, overestimating stability among a sample that is more interested in their political attitudes, and committed to completing surveys over time). Analyses by Sibley et al. (2017) indicate, relative to another major election polling source, the NZAVS does well at tracking party support over time. Yet the study does tend to over-estimate Green party support, and slightly underestimate Labour Party support. However, around 7% of the sample for our 5-year stability estimates provided the minimum rating of political identity centrality, with a mean at about the mid-point of the scale, suggesting our sample is not notably biased towards those with an interest in politics. In this sense, because the NZAVS is an omnibus study that assesses a range of psychological variables, it may be less prone to over-sampling the politically engaged than a survey with explicit ties to political beliefs. Application of sample weighting based on gender, ethnicity, and region also had minimal effects on our standardized stability estimates (with some slightly decreasing, and others slightly increasing).

Table 1. Means, standard deviations, and correlations across waves.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. Gender (2012)	-														
2. Age (2012)	.137*	-													
3. Political identity centrality (2014)	.053*	.095*	-												
4. National support (2012)	.063*	.098*	-.124*	-											
5. Labour support (2012)	-.105*	-.096*	.100*	-.568*	-										
6. Green support (2012)	-.175*	-.179*	.107*	-.474*	.516*	-									
7. NZF support (2012)	.031	.052*	-.044*	-.126*	.163*	.092*	-								
8. National support (2014)	.078*	.113*	-.123*	**.842***	-.579*	-.507*	-.129*	-							
9. Labour support (2014)	-.099*	-.061*	.120*	-.569*	**.721***	.464*	.101*	-.532*	-						
10. Green support (2014)	-.182*	-.203*	.129*	-.515*	.494*	**.757***	.015	-.493*	.536*	-					
11. NZF support (2014)	.026	.076*	-.017	-.120*	.112*	.022	**.583***	-.125*	.193*	.111*	-				
12. National support (2016)	.077*	.115*	-.128*	**.780***	-.560*	-.523*	-.097*	**.822***	-.542*	-.541*	-.102*	-			
13. Labour support (2016)	-.130*	-.064*	.077*	-.555*	**.658***	.506*	.073*	-.561*	**.680***	.547*	.106*	-.620*	-		
14. Green support (2016)	-.173*	-.180*	.116*	-.485*	.485*	**.710***	-.012	-.502*	.487*	**.740***	-.008	-.523*	.611*	-	
15. NZF support (2016)	.018	.085*	-.002	-.207*	.191*	.090*	**.496***	-.198*	.178*	.099*	**.563***	-.196*	.256*	.148*	-
Mean	.37	49.10	4.10	4.06	4.13	4.11	2.71	4.12	4.09	4.16	3.15	3.99	4.58	4.26	3.10
SD	.48	15.03	1.71	2.00	1.74	1.84	1.59	2.06	1.65	1.82	1.56	1.99	1.75	1.84	1.58

Note. * $p < .001$. Test-retest correlations between party support measures are boldface.

Table 2. Regression model assessing moderators of National Party support stability from 2012 – 2014 and 2012 – 2017.

	National support (2014)				National support (2017)			
	β	*B*	*SE*	*p*	*β*	*B*	*SE*	*p*
National support (2012)	**.840****	.878	.009	< .001	**.770****	.794	.012	< .001
Gender	.002	.008	.027	.763	.005	.020	.035	.564
Age	.044**	.062	.009	< .001	.022	.030	.023	.198
Age^2	-.019*	-.015	.005	.005	-.007	-.006	.007	.423
Age^3	-	-	-	-	.033*	.008	.004	.045
Political identity centrality	-.019*	-.022	.008	.006	-.037**	-.045	.011	< .001
Support × gender	-.009	-.016	.012	.177	.002	.003	.016	.867
Support × age	.012*	.009	.004	.033	.058**	.042	.011	< .001
Support × age^2	-.018*	-.005	.003	.029	-.029*	-.009	.003	.006
Support × age^3	-	-	-	-	-.030*	-.004	.002	.042
Support x identity	.051**	.031	.004	< .001	.076**	.046	.005	< .001

Note: * *p* < .05, ** *p* < .001. Standard errors are of the standardized coefficients. Focal standardized 2-year and 5-year test-retest stability estimates are boldface.
N(2014) = 7,288, R^2 = .72, *N*(2017) = 5,481, R^2 = .63.

Table 3. Regression model assessing moderators of Labour Party support stability from 2012 – 2014 and 2012 – 2017.

	Labour support (2014)				Labour support (2017)			
	β	*B*	*SE*	*p*	*β*	*B*	*SE*	*p*
Labour support (2012)	**.715****	.691	.013	< .001	**.653****	.682	.017	< .001
Gender	-.018*	-.061	.028	.027	-.046*	-.170	.037	< .001
Age	-.016	-.018	.010	.067	-.013	-.016	.013	.223
Age^2	.003	.002	.005	.698	.009	.007	.007	.376
Political identity centrality	.051**	.049	.008	< .001	.031*	.033	.011	.004
Support × gender	-.005	-.008	.017	.625	.004	.007	.022	.752
Support × age	.019*	.013	.006	.030	.040**	.029	.008	< .001
Support × age^2	-.023*	-.007	.003	.049	-.042*	-.013	.005	.003
Support × identity	.080**	.044	.005	< .001	.082**	.050	.007	< .001

Note: * *p* < .05, ** *p* < .001. Standard errors are of the standardized coefficients. Focal standardized 2-year and 5-year test-retest stability estimates are boldface.
N(2014) = 7,261, R^2 = .53, *N*(2017) = 5,483, R^2 = .45.

Table 4. Regression model assessing moderators of Green party support stability from 2012 – 2014 and 2012 – 2017.

	Green support (2014)				Green support (2017)			
	β	*B*	*SE*	*p*	*β*	*B*	*SE*	*p*
Green support (2012)	**.712****	.709	.010	< .001	**.676****	.692	.012	< .001
Gender	-.041*	-.155	.029	< .001	-.037**	-.144	.037	< .001
Age	-.073**	-.092	.010	< .001	-.054**	-.071	.013	< .001
Political identity centrality	.042**	.044	.008	< .001	.077**	.086	.011	< .001
Support × gender	.017	.026	.015	.073	.001	.001	.019	.945
Support × age	-.008	-.005	.005	.297	.003	.002	.006	.705
Support × identity	.076**	.042	.004	< .001	.067**	.040	.006	< .001

Note: * *p* < .05, ** *p* < .001. Standard errors are of the standardized coefficients. Focal standardized 2-year and 5-year test-retest stability estimates are boldface.
N(2014) = 7,250, R^2 = .59, *N*(2017) = 5,461, R^2 = .52.

Table 5. Regression model assessing moderators of New Zealand First support stability from 2012 – 2014 and 2012 – 2017.

	NZ First support (2014)				NZ First support (2017)			
	β	*B*	*SE*	*p*	*β*	*B*	*SE*	*p*
NZ First support (2012)	**.579****	.582	.016	< .001	**.508****	.534	.020	< .001
Gender	.005	.017	.031	.574	.003	.011	.039	.776
Age	.057*	.061	.020	.002	.062*	.070	.026	.006
Age^2	.015	.009	.006	.140	-.038*	-.024	.008	.002
Age^3	-.028	-.005	.003	.138	-.030	-.006	.004	.192
Political identity centrality	-.006	-.006	.009	.525	.026*	.025	.012	.035
Support × gender	-.001	-.001	.021	.945	-.010	-.017	.027	.522
Support × age	-.019	-.013	.013	.306	-.006	-.004	.018	.813
Support × age^2	-.005	-.001	.004	.721	-.013	-.005	.005	.373
Support × age^3	.043*	.005	.002	.023	.052*	.007	.003	.025
Support × identity	.027*	.016	.006	.011	.004	.003	.008	.730

Note: * $p < .05$, ** $p < .001$. Standard errors are of the standardized coefficients. Focal standardized 2-year and 5-year test-retest stability estimates are boldface.
N(2014) = 7,222, R^2 = .34, N(2017) = 5,449, R^2 = .25.

Table 6. Standardized stability estimates solved at high and low levels of political identity centrality.

	National support		Labour support		Green support		NZ First support	
	2014	2017	2014	2017	2014	2017	2014	2017[1]
High centrality	.890	.844	.792	.733	.785	.742	.606	.512
Low centrality	.790	.695	.639	.573	.639	.611	.553	.504

Note: Party support stability estimates solved at +/- 1 SD from the mean, estimated based on the standardized parameters presented in Tables 2-5.
[1] Political identity centrality was not a significant moderator of the 5-year stability of NZF support.

Notably, the stability estimates of party support were significantly moderated by political identity centrality, except for the stability of NZ First support over the 5-year test-retest period. Table 6 displays the stability estimates for party support across each period solved at high and low levels of identity centrality based on the models presented in Tables 2-5. Over the two-year period, moving from low to high identity centrality produced a .05 - .15 increase in party support stability across the parties. Over the five-year period, stability estimates of party support among those with high identity centrality were between .13 to .16 units higher than at low levels of centrality. These results are consistent with the expressive approach of partisanship, which emphasizes stability as a product of strong social identities.

Results also revealed age differences in the stability of party support for most parties. Figure 1 displays the trajectory of 5-year support stability across the lifespan for each party, with different age functions observed for each party. In terms of the major parties, Labour Party support generally increased in stability with older age, but National Party support decreased in younger age, before increasing again between roughly ages 36 – 60, and decreasing thereafter. Although we expected party support for these parties to generally increase with age, the cubic function of National Party support stability was only marginally significant given our sample size (i.e., $p = .042$) for the 5-year period, and the function was more consistent with expectations over the 2-year period (i.e., stability increasing with age, albeit at a decreasing rate). In terms of the minor parties, age did not moderate the stability of Green party support, whereas a cubic function was identified for NZ First support. Specifically, the stability of party support for NZ First was low amongst the young, increased to a plateau in mid-life, and increased further amongst the oldest in our sample. This pattern seems consistent with NZ First's role in New Zealand politics as a relatively less salient party that tends to advocate for senior citizens' rights.

Markov Models of vote stability and switching

To examine whether test-retest stability estimates of party support aligned with the stability of voting behaviour, we utilized Markov Modelling to investigate the patterns of stability and change in respondents' reported party vote across three national elections (2011 and 2014, both of which the National Party was elected into government, and 2017, where the Labour Party formed a government). For these analyses, we categorized party votes into four major categories: 'National Party', 'Labour Party', 'minor party', or 'no vote' in the election. Minor party votes were predominantly for the Green party (11.8-16.8% of the sample across the years) relative to NZ

Table 7. Results of the Markov Model estimating the effects of National and Labour Party support on the log-odds of changing vs. retaining National or Labour Party vote from 2011 – 2014.

	2014 vote					
Predictors among National Party voters in 2011	Labour vote vs. National vote		Minor vote vs. National vote		No vote vs. National vote	
	b	*SE*	*b*	*SE*	*b*	*SE*
National support	-1.128**	.107	-.666**	.070	-.262*	.115
Labour support	1.156**	.124	.110	.071	.274*	.102
Predictors among Labour Party voters in 2011	National vote vs. Labour vote		Minor vote vs. Labour vote		No vote vs. Labour vote	
	b	*SE*	*b*	*SE*	*b*	*SE*
National support	.934**	.106	.090	.059	.328*	.126
Labour support	-.449**	.109	-.284**	.066	-.386*	.158

Note: N = 5,449. * $p < .05$ ** $p < .001$

Table 8. Results of Markov Model estimating the effects of National and Labour Party support on the log-odds of changing vs. retaining National or Labour Party vote from 2014 – 2017.

	2017 vote					
Predictors among National Party voters in 2014	Labour vote vs. National vote		Minor vote vs. National vote		No vote vs. National vote	
	b	*SE*	*b*	*SE*	*b*	*SE*
National support	-.808**	.053	-.640**	.059	-.453**	.097
Labour support	.591**	.048	.227**	.057	.105*	.073
Predictors among Labour Party voters in 2014	National vote vs. Labour vote		Minor vote vs. Labour vote		No vote vs. Labour vote	
	b	*SE*	*b*	*SE*	*b*	*SE*
National support	.627**	.085	.039	.054	.289*	.085
Labour support	-.629**	.103	-.305**	.059	-.172	.136

Note: N = 9,845. * $p < .05$ ** $p < .001$

First (3.3-5.1%) or other party (5.1-6.3%) votes. National and Labour Party voters comprised 35-40.5% and 20.8-33.9% of the sample across the years, respectively. Those explicitly indicating that they did not vote made up 5.2-7.9% of valid responses across the years, while people who indicated that they were unsure or did not report who they voted for (3.9-6.1% of valid responses) were excluded from these analyses.

Figure 2 displays the results of a Markov Model estimating vote transitions across the 2011, 2014, and 2017 elections without covariates. Consistent with the results for the stability of party support, voting for the National Party tended to be most stable over time. The stability of voting for the Labour Party tended to be somewhat lower across the 2011 – 2014 elections, but was highly stable from 2014 – 2017. Across both election cycles, the probability of switching votes between these two major parties was very low, ranging from .04 (Labour voters becoming National voters in 2017) and .12 (National voters becoming Labour voters in 2017). Consistent with past research showing that some New Zealanders solely support Labour, whereas others support both Labour and the Green Party (Greaves et al, 2014), the highest probabilities for vote switching occurred to-and-from the Labour Party and minor parties (predominantly Green Party voters).

New Zealanders can strongly support multiple parties, particularly Labour and the Greens (Greaves et al., 2015), and may also engage in strategic voting based on their perceptions of party performance (Bowler et al., 2010). Thus, we ran two additional models that included support for the National Party and Labour Party as covariates (see Tables 7 and 8). These models reveal whether a) party support is distinct from, and not simply a restatement of,

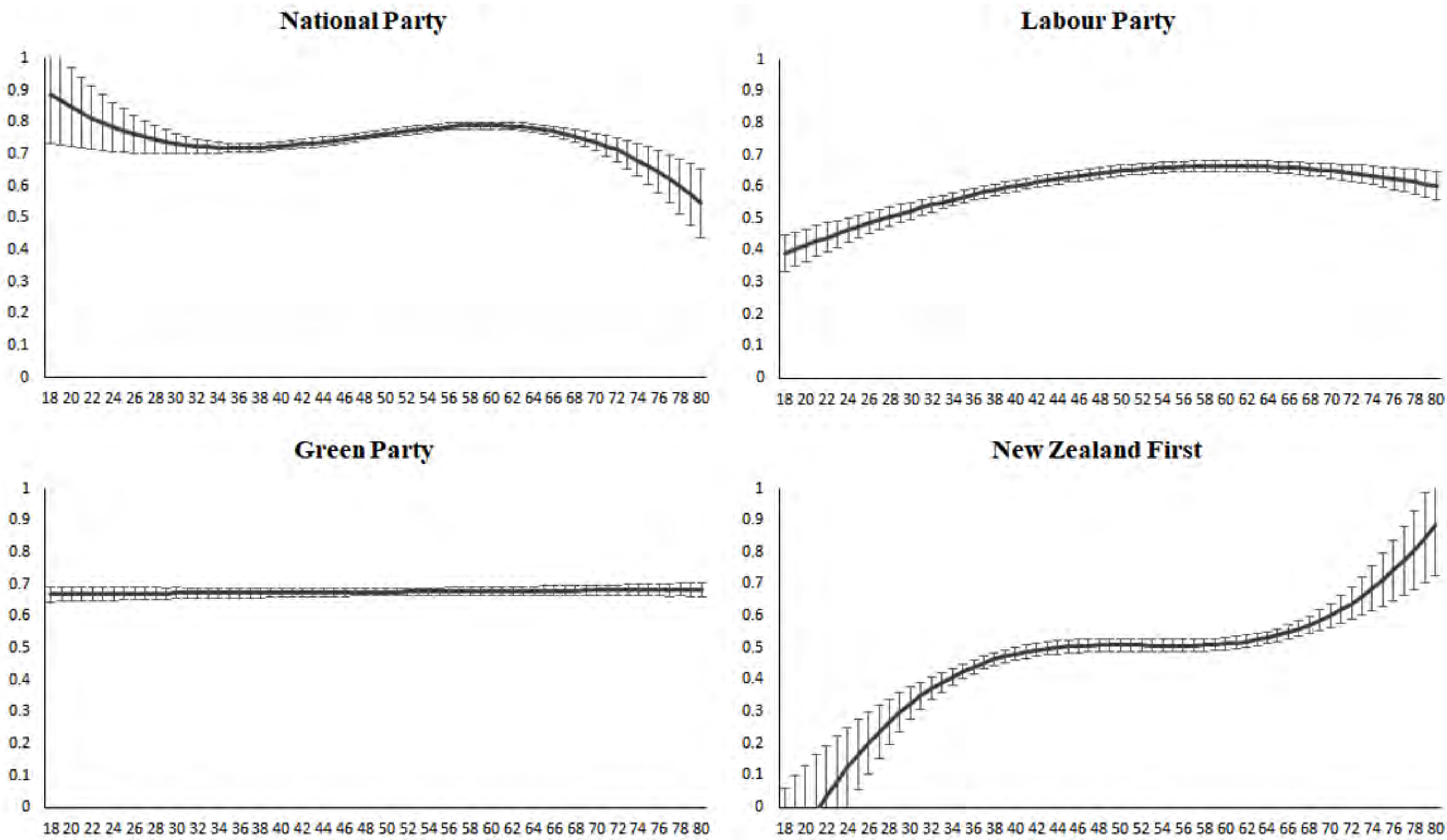


Figure 1. Standardized 5-year stability estimates of party support (y-axis) as a function of age (from 18 – 80 years; x-axis). Error bars represent standard errors of the estimates.

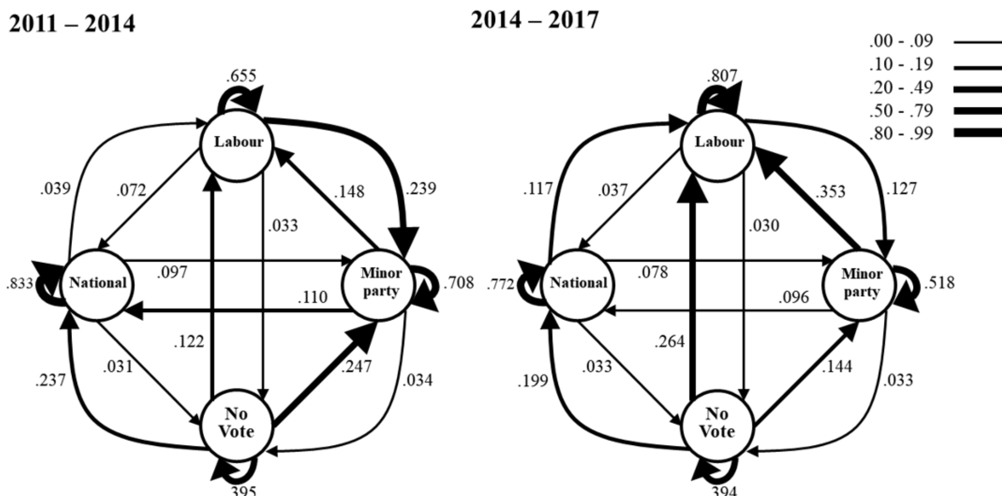


Figure 2. Markov Models estimating the transition probabilities of maintaining the same vote, and switching votes between major and minor parties, and non-voting, across the 2011 – 2014 (left) and 2014 – 2017 (right) New Zealand general elections (N = 6,232).

Strong National party supporters

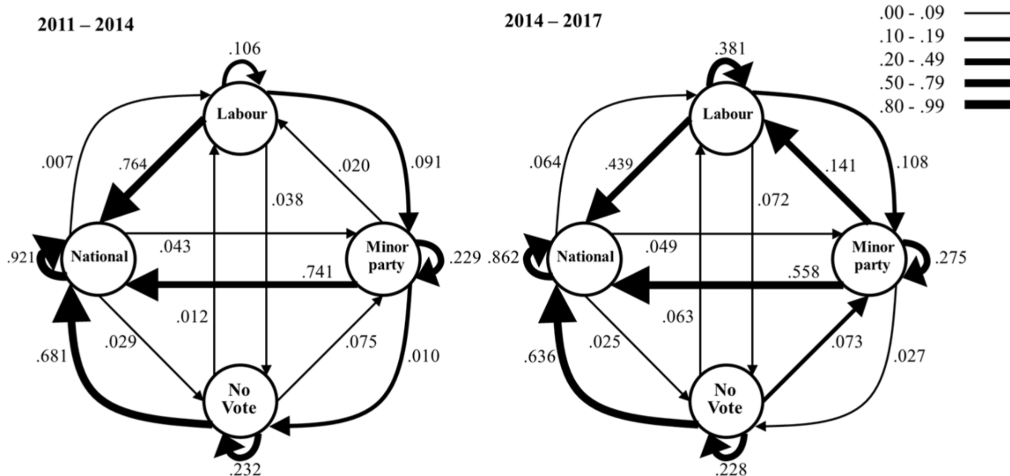


Figure 3. Transition probabilities of maintaining vs. switching party vote across elections among those maximally supportive of the National Party (Labour Party support held at mean levels; $N_{2011-2014}$ = 5,449. $N_{2014-2017}$ = 9,845).

Strong Labour party supporters

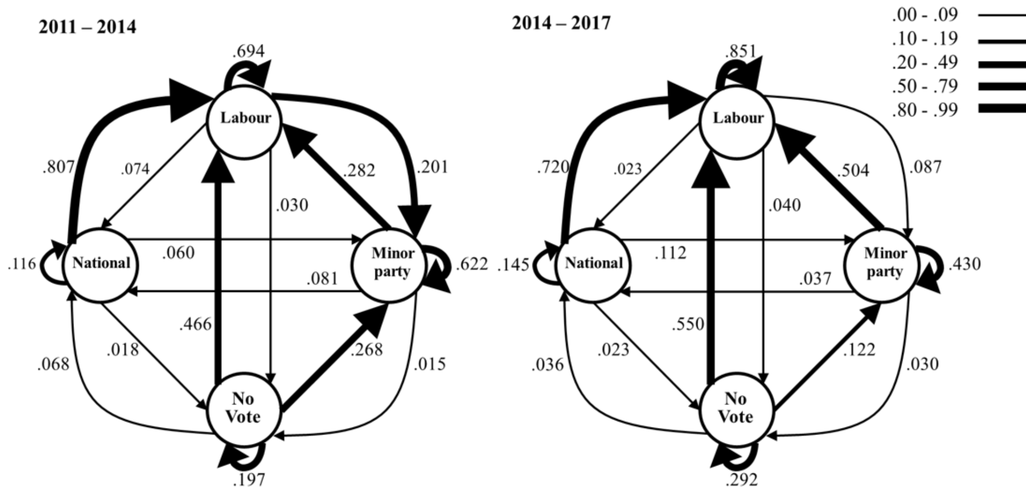


Figure 4. Transition probabilities of maintaining vs. switching party vote across elections among those maximally supportive of The Labour Party (National Party support held at mean levels; $N_{2011-2014}$ = 5,449. $N_{2014-2017}$ = 9,845).

voting, and b) whether participants vote strategically, where supporters of a given party (Labour) may nonetheless demonstrate high probabilities of voting for minor parties (e.g., the Greens) when Labour performs poorly (or vice versa).

Consistent with our expectations, party support significantly predicted future voting behaviour, even when accounting for prior party vote. The probabilities of stable voting vs. vote switching across elections are displayed in Figure 3 (solved at maximal support for the National Party and mean support for the Labour Party) and Figure 4 (maximal support for the Labour Party and mean support for the National Party). Compared to the model without party support covariates, these models demonstrate that consistent party voting across elections was even more probable among voters strongly supportive of the National or Labour Party. Moreover, people who were strongly supportive of National or Labour who (for whatever reason) did not vote for these parties in a prior election had very high probabilities of returning their vote to their strongly supported party in the following election.

Findings were, however, slightly more nuanced for strong Labour supporters. Strong Labour Party supporters who voted for Labour in 2011 had a reasonably high probability (.201) of voting for a minor party (e.g., the Greens) in 2014, whereas strong Labour supporters who voted for a minor party were likely to maintain a minor party vote in 2014. Yet, in 2017, when the Labour Party gained momentum under Jacinda Ardern's leadership, strong Labour supporters who voted for a minor party in 2014 had a .50 probability of voting for Labour in 2017, while strong Labour supporters who voted Labour in 2014 were only about one third as likely to switch to a minor party vote in 2017 compared to the same probability of switching from 2011 – 2014. These results are consistent with the presence of voters who are highly supportive of *both* the Labour and Green parties, and seem to reflect an element of strategic voting (such that these voters will switch their votes between the two parties, depending on perceived party performance).

DISCUSSION

Stability is considered a crucial aspect of partisanship that indicates a commitment to a party, rather than a fleeting judgement. It implies citizens remain committed even in the face of undesirable performance or policies, yet its presence in multiparty systems has remained in question. In this study, we assessed the stability of party support and voting as attitudinal and behavioural indicators of partisanship in New Zealand, where stability has not been systematically analysed in recent years. By examining party support, we also provide useful information on an attitudinal measure that can accommodate the presence of multiple party commitments, as well as both 'negative' and 'positive' partisanship in multiparty systems. We used a longitudinal national probability sample of New Zealand adults to assess the 2 and 5-year stability of party support (attitudinal partisanship), and the stability of party voting (behavioural partisanship) between three national elections from 2011 – 2017. The results demonstrated high levels of stability in New Zealanders' support for both major and minor parties over 2-year (.58 - .84) and 5-year (.51 - .77) test-retest periods. Moreover, stability estimates were generally higher among participants who indicated that their political beliefs were important to their sense of self (.61 - .89 for the 2-year period, and .51 - .85 for the 5-year period). Finally, generally high levels of stable party voting between national elections were observed (consistent party vote probabilities of .66 - .83).

To place these estimates within context, the stability of support for major parties in particular does not stray far from the 2-year test-retest estimates demonstrated by Big-Six personality traits, which are generally considered highly stable over time (i.e., .73 - .92; Milojev & Sibley, 2014). In contrast, the stability of party support generally exceeded the 3-year test-retest estimates of Schwartz values (.55 - .60; see Milfont, Milojev, & Sibley, 2016; note that these papers used the same modelling method used here). These results therefore demonstrate that party support, at least in New Zealand, reflects stable and meaningful attitudes toward political parties. Support also represents more than a restatement of voting behaviour, as party support predicts future voting behaviour while adjusting for prior vote. This is particularly consequential when considering the sway political parties can have on citizens' attitudes. Highly stable support for parties, particularly among those whose identities are invested in politics, might indicate greater susceptibility to follow the party position on political issues. Such effects have already been demonstrated in New Zealand (see Satherley et al., 2018; Satherley, Osborne, & Sibley, 2019), thus reinforcing the view that partisanship can have detrimental effects, beyond the U.S.

Our findings also reveal nuances in the stability of partisan attitudes in multiparty systems in general. Consistent with our hypothesis and past research (e.g. Kuhn, 2009; Richardson, 1991), support for major parties was generally more stable than support for minor parties. Yet, our results also suggest the presence of multiple party commitments (e.g., Greaves et al., 2015), which may lead to lower attitudinal, and particularly behavioural (i.e. voting), partisan stability in multiparty systems. Our results demonstrate that strong Labour supporters maintained reasonably high probabilities of switching their vote to a minor party (most likely the Greens) in 2014, when Labour was performing relatively poorly, compared to 2017, when Labour's popularity soared under Jacinda Ardern's leadership. This is broadly consistent with Bowler et al.'s (2010) analyses of strategic voting in the 2002 New Zealand election, indicating New Zealanders were more likely to vote for their second-choice party if they believed that their preferred party would not win. For example, Labour supporters may have switched to a Green vote to instead ensure the party stayed above the 5% parliamentary threshold when it seemed Labour was unlikely to win, but returned their vote to Labour when the prospect of victory seemed greater (or vice versa). Overall, these findings suggest behavioural measures of partisanship, such as party voting, may be inherently less stable in multiparty systems.

Our party-by-party approach to examining stability in support and voting may also explain the relatively higher image of stability conveyed here compared to Vowles (2016) who found both relatively low party identification in 2014 (40%) and relatively high vote volatility (just

under 40% of New Zealand Election Study participants switched their vote across the 2008 – 2014 elections). Our results show that stability in voting is much higher for major parties to compared to minor parties and non-voting. Thus, composite measures of volatility across parties (like those presented by Vowles, 2016) will produce an average estimate. Moreover, data from Vowles (2018) shows that the effective number of parties in New Zealand has been relatively stable since 2005, and much lower than the period between 1996-2002. For example, the effective number of parties was at 2.9 and 3.3 in 2017 and 2014, respectively, compared to a peak of 4.4 in 1996. Thus, a smaller and more stable number of represented parties may have given way over time to more stable support for those parties in recent years. Finally, the low probability for people to switch votes between the Labour and National parties appears consistent with the ideological polarization of parties in New Zealand (Dalton, 2008). Despite the two major parties being centre-left and centre-right, New Zealanders still appear to perceive large enough differences between the parties to consider them quite distinct.

Finally, it should be noted that our analyses are impacted by the political context at the time. Although our analyses are longitudinal, they remain limited to a relatively small, stable timeframe (2011 – 2017) in New Zealand politics. As such, indices of stability may change depending on the political context of the time. This is hinted at in our analyses, as the stability of support and voting for Labour, a major party who underwent a number of leadership changes and tended to poll poorly, was lower than for National, a successful party that maintained reasonably steady electoral support over the period. Data from the 2020 election would provide a strong test of the impact of political conditions on the stability of support and voting for National, as their vote share dropped by 18.9 percentage points from the previous election, whereas the Labour Party's vote share soared as a result of their internationally-praised response to the COVID-19 pandemic. Future analyses could examine specifically where the votes of prior National Party voters shifted (e.g., to Labour, the ACT party, or to abstention), and whether the sharp drop in votes for the National Party also corresponded with a drop in ratings of support for the party.

References

Aimer, P. (1989). Travelling together: Party identification and voting in the New Zealand general election of 1987. *Electoral Studies*, *8*(2), 131–142. https://doi.org/10.1016/0261-3794(89)90030-9

Bartels, L. M. (2000). Partisanship and voting behavior, 1952-1996. *American Journal of Political Science*, 35-50. doi: 10.2307/2669291

Bartle, J., & Bellucci, P. (2009). Introduction: Partisanship, social identity and individual attitudes. In J. Bartle and P. Belluci (Eds.). *Political Parties and Partisanship* (pp. 1-25). New York: Routledge.

Bolsen, T., Druckman, J. N., & Cook, F. L. (2014). The influence of partisan motivated reasoning on public opinion. *Political Behavior*, *36*(2), 235-262. doi: 10.1007/s11109-013-9238-0

Bowler, S., Karp, J. A., & Donovan, T. (2010). Strategic coalition voting: evidence from New Zealand. *Electoral Studies*, *29*(3), 350-357. doi: 10.1016/j.electstud.2010.03.001

Campbell, A., Converse, P. E., Miller, W., & Stokes, D. (1960). *The American voter*. Chicago, IL: John Wiley & Sons.

Carsey, T. M., & Layman, G. C. (2006). Changing sides or changing minds? Party identification and policy preferences in the American electorate. *American Journal of Political Science, 50*, 464-477. doi: 10.1111/j.1540-5907.2006.00196.x

Caruana, N. J., McGregor, R. M., & Stephenson, L. B. (2015). The power of the dark side: Negative partisanship and political behaviour in Canada. *Canadian Journal of Political Science/Revue canadienne de science politique, 48*, 771-789. doi: 10.1017/S0008423914000882

Dalton, R. J. (2008). The quantity and the quality of party systems: Party system polarization, its measurement, and its consequences. *Comparative Political Studies*, *41*(7), 899-920. https://doi.org/10.1177/0010414008315860

Dalton, R. J., & Weldon, S. (2007). Partisanship and Party System Institutionalization. *Party Politics*, *13*(2), 179–196. https://doi.org/10.1177/1354068807073856

Dassonneville, R., Blais, A., & Dejaeghere, Y. (2015). Staying with the party, switching or exiting? A comparative analysis of determinants of party switching and abstaining. *Journal of Elections, Public Opinion and Parties, 25*(3), 387-405. https://doi.org/10.1080/17457289.2015.1016528

Fiorina, M. P. (1981). *Retrospective voting in American national elections*. New Haven: Yale University Press.

Franklin, C. H., & Jackson, J. E. (1983). The dynamics of party identification. *American political science review*, *77*(4), 957-973. doi: 10.2307/1957569

Goren, P., Federico, C. M., & Kittilson, M. C. (2009). Source cues, partisan identities, and political value expression. *American Journal of Political Science, 53*, 805-820. doi: 10.1111/j.1540-5907.2009.00402.x

Greaves, L. M., Osborne, D., Sengupta, N. K., Milojev, P., & Sibley, C. G. (2014). Politics and post-colonial ideology: Historical negation and symbolic exclusion predict political party preference. *New Zealand Journal of Psychology*, *43*(3), 39-54.

Greaves, L. M., Osborne, D., & Sibley, C. G. (2015). Profiling the Fence-Sitters in New Zealand Elections: A Latent Profile Model of Political Voting Blocs. *New Zealand Journal of Psychology, 44*(2), 43-56.

Green, D. P., & Palmquist, B. (1994). How stable is party identification? *Political behavior, 16*, 437-466. doi: 10.1007/BF01498826

Green, D. P., Palmquist, B., & Schickler, E. (2002). *Partisan hearts and minds: Political parties and the social identity of voters*. New Haven, CT: Yale University Press.

Huddy, L., Bankert, A., & Davies, C. (2018). Expressive Versus Instrumental Partisanship in Multiparty European Systems. *Political Psychology, 39*, S1, 173-199. doi: 10.1111/pops.12482

Johnston, R. (2006). Party identification: Unmoved mover or sum of preferences? *Annual Review of Political Science, 9*, 329-351. doi: 10.1146/annurev.polisci.9.062404.170523

Kuhn, U. (2009). Stability and change in party preference. *Swiss Political Science Review, 15*, 463-494. doi: 10.1002/j.1662-6370.2009.tb00142.x

Lamare, J. W. (1984). Party Identification and Voting Behaviour in New Zealand. *Political Science*, *36*(1), 1–9. https://doi.org/10.1177/003231878403600101

Leithner, C. (1997). Of Time and Partisan Stability Revisited: Australia and New Zealand 1905-90. *American Journal of Political Science*, *41*(4), 1104–1127. https://doi.org/10.2307/2960483
Mayer, S. J. (2017). How negative partisanship affects voting behavior in Europe: Evidence from an analysis of 17 European multi-party systems with proportional voting. *Research & Politics, 4*(1). doi: 10.1177/2053168016686636
Milfont, T. L., Milojev, P., & Sibley, C. G. (2016). Values stability and change in adulthood: A 3-year longitudinal study of rank-order stability and mean-level differences. *Personality and Social Psychology Bulletin, 42*, 572-588. doi: 10.1177/0146167216639245
Milojev, P., & Sibley, C. G. (2014). The stability of adult personality varies across age: Evidence from a two-year longitudinal sample of adult New Zealanders. *Journal of Research in Personality, 51*, 29-37. doi: 10.1016/j.jrp.2014.04.005
Petrocik, J. R. (2009). Measuring party support: Leaners are not independents. *Electoral Studies*, *28*(4), 562-572. doi: 10.1016/j.electstud.2009.05.022
Reijlan, A. (2020). Fear and loathing across party lines' (also) in Europe: Affective polarisation in European party systems. *European Journal of Political Research, 59*. https://doi.org/10.1111/1475-6765.12351
Richardson, B. M. (1991). European party loyalties revisited. *American Political Science Review, 85*, 751-775. doi: 10.2307/1963849
Rose, R., & Mishler, W. (1998). Negative and positive party identification in post-communist countries. *Electoral Studies*, *17*(2), 217-234. doi: 10.1016/S0261-3794(98)00016-X
Satherley, N., Osborne, D., & Sibley, C. G. (2019). Who Is for (or Against) the National Flag? Ideological and Identity-Based Motivators of Attitudes. *Analyses of Social Issues and Public Policy, 49*. 407-428. https://doi.org/10.1111/asap.12169
Satherley, N., Yogeeswaran, K., Osborne, D., & Sibley, C. G. (2018). If they say 'yes', we say 'no': partisan cues increase polarisation over national symbols. *Psychological Science*, *29*, 1996-2009. doi: 10.1177/0956797618805420.
Schickler, E., & Green, D. P. (1997). The stability of party identification in western democracies: results from eight panel surveys. *Comparative Political Studies, 30*, 450-483. doi: 10.1177/0010414097030004003
Schmitt-Beck, R., Weick, S., & Christoph, B. (2006). Shaky attachments: Individual-level stability and change of partisanship among West German voters, 1984–2001. *European Journal of Political Research, 45*, 581-608. doi: 10.1111/j.1475-6765.2006.00310.x
Sears, D. O., & Funk, C. L. (1999). Evidence of the long-term persistence of adults' political predispositions. *The Journal of Politics, 61*, 1-28. doi: 10.2307/2647773
Sibley, C. G., Robertson, A., Osborne, D., Huang, Y., Milojev, P., Greaves, L. M., ... & Barlow, F. K. (2017). Bias and tracking accuracy in voting projections using the New Zealand attitudes and values study. *Political Science*, *69*(1), 16-34.
Stoker, L., & Jennings, M. K. (2008). Of time and the development of partisan polarization. *American Journal of Political Science, 52*, 619-635. doi: 10.1111/j.1540-5907.2008.00333.x
Thomassen, J., & Rosema, M. (2009). Party identification revisited. In J. Bartle and P. Belluci (Eds.). *Political Parties and Partisanship* (pp. 42-59). New York: Routledge.
Vowles (2005). New Zealand: The Consolidation of Reform? In M. Gallagher & P. Mitchell (Eds.) *The politics of electoral systems* (pp. 295–312). New York: Oxford University Press.
Vowles, J. (2016). Two Cheers for MMP: A Twenty-Year Review. Plenary Address to the New Zealand Political Studies Association Annual Conference, University of Waikato.
Vowles, J. (2018). Surprise, surprise: the New Zealand general election of 2017. *Kōtuitui: New Zealand Journal of Social Sciences Online*, *13*(2), 147-160. doi: 10.1080/1177083X.2018.1443472
Weisberg, H. F. (1980). A multidimensional conceptualization of party identification. *Political Behavior*, *2*(1), 33-60. doi: 10.1007/BF00989755

Corresponding Author
Nicole Satherley
School of Psychology,
University of Auckland,
Private Bag 92019,
Auckland 1142,
New Zealand
Email: nsat639@aucklanduni.ac.nz

Acknowledgements
This research forms part of the PhD thesis of Nicole Satherley. Collection of the NZAVS data analysed in this paper was supported by a grant from the Templeton Religion Trust (TRT0196) awarded to Chris Sibley. Nicole Satherley was supported by a University of Auckland Doctoral Scholarship. Syntax for the analyses are available on the NZAVS website: www.psych.auckland.ac.nz/uoa/NZAVS. A copy of the data is available from Chris Sibley upon request for the purpose of replicating the analyses reported here

He Piki Raukura: Assessing Ao Māori developmental constructs – Part I: Reliability of novel strengths-based measures among preschool Māori children

Aroaro Tamati[1,2], Gareth J. Treharne[2], Reremoana Theodore[2], Mihi Ratima[1,2], Erana Hond-Flavell[1,2], Will Edwards[1,2], Ruakere Hond[1,2], Richie Poulton[2]
[1] Te Pou Tiringa Incorporated, New Plymouth, Taranaki, Aotearoa/New Zealand
[2] National Centre for Lifecourse Research, Department of Psychology, University of Otago, Aotearoa/New Zealand

This is the first of two papers describing the creation of measurement tools for four Māori constructs of positive child behaviour – tuakiri (secure local Māori identity); whānauranga (acting as a member of a whānau); manawaroa (persisting despite difficulty); and piripono (having integrity, commitment and responsibility). This paper describes the psychometric properties of these new measures. Parents and teachers completed questionnaires on 28 children aged 0-5 years five times over 10 months in a Māori-medium early years setting, and video observations were made. Ratings of the videos showed good inter-rater reliability. All questionnaire measures had good internal consistency. Associations of questionnaires with rated observations varied at some timepoints suggesting a need to include both in ongoing research. This study provides initial evidence about reliability of our novel Māori measurement tools for assessment of preschool Māori children.

Keywords: *Māori Indigenous psychology, Māori constructs, Māori child behaviour measures, Māori child development, reliability, validity*

Introduction

In Aotearoa, a number of measurement tools are used to assess and evaluate young children's development and behaviour. These assessments are commonly carried out by psychologists in research or practice settings, through government programmes such as Plunket's Well Child Tamariki Ora, or in early childhood settings and primary schools (Pannekoek & D'Souza, 2018). Assessments typically involve measures created by non-Māori researchers such as the Strengths and Difficulties Questionnaire or the Social Competence Scale which assess prosocial behaviours and conduct problems (Corrigan, 2002; Goodman, 1997). Tamariki Māori are automatically included in assessments using these measurement approaches (Morton et al., 2017; Peterson et al., 2018), despite the tools being created by non-Māori researchers and practitioners. What this means is that measurement tools often do not take into account Māori cultural priorities and other Indigenous factors, such as the child's cultural context, language, and whānau, hapū and iwi connections (Bishop, Berryman, Cavanagh, & Teddy, 2009; Durie, 2006; G. H. Smith, 2003b).

There is a need to create child behaviour measurement tools, that are both strengths-based and grounded within Indigenous Māori worldviews. There is strong evidence in Aotearoa that interventions in the early years (i.e., 0-5 years) lead to improved life outcomes (Fergusson, Horwood, Ridder, & Grant, 2005; Horwood, Gray, & Fergusson, 2011; Sturrock, Gray, Fergusson, Horwood, & Smits, 2014). Running parallel with this are the growing number of initiatives, programmes and approaches by Māori communities or groups, which are increasingly recognised as alternative ways to work with tamariki Māori, whānau and community to improve life outcomes for Māori (Durie, Cooper, Grennell, Snively, & Tuaine, 2010; Hond, 2013; King & Turia, 2002; Mane, 2009; Royal Tangaere, 2012). These include Māori language, health and education initiatives. However, to date, there have been few evaluation studies conducted in these early life kaupapa Māori community initiatives. Moreover, deficit theories have tended to dominate the way outcomes of interest to Māori have been analysed (Blank, Houkamau, & Kingi, 2016; Pihama, 2012), further emphasising the need for the development and validation of strengths-based Māori measurement tools to assess the development of tamariki Māori and the outcomes of Māori programmes and interventions.

While there are many studies that have examined the development of measures of young children's behaviour (D'Souza, Waldie, Peterson, Underwood, & Morton, 2017; Goodman, 2001; Ponitz et al., 2008; Rothbart, Ahadi, Hershey, & Fisher, 2001), there are few psychological measures that have been developed from within Indigenous worldviews. Those that have been developed have focussed on Indigenous youth and adults, such as measures of Māori identity in adulthood (Houkamau & Sibley, 2010a; Houkamau & Sibley, 2010b; Palmer, 2004; Sibley & Houkamau, 2013). Examples from overseas include measures of protective factors in Alaskan youth engaged in alcohol abuse and for suicide prevention (Allen et al., 2014); growth and empowerment in Indigenous Australians (Haswell et al., 2010); cognitive assessment of rural-based middle-aged Indigenous Australians (LoGiudice et al., 2011); and the

emotional intelligence of Indigenous adults in Pakistan (Batool & Khalid, 2011).

Measures that have been created from within Indigenous worldviews or adapted through application of an Indigenous cultural lens have been applied in Indigenous-specific research. Examples include Indigenous language assessment in children and Māori parenting interventions (Housman, Dameg, Kobashigawa, & Brown, 2011; Keown, Sanders, Franke, & Shepherd, 2018). In the development of the measures, these studies used culturally-grounded approaches including collaboration, community involvement and contribution (Batool & Khalid, 2011; Keown et al., 2018); iterative processes of dialogue and workshopping (Allen et al., 2014); participation of Indigenous experts on health and education (Schlesinger, Ober, McCarthy, Watson, & Seinen, 2007); and the initial generation of items from within an Indigenous language context (Batool & Khalid, 2011; Housman et al., 2011).

Cultural psychometrics

The measures used in the assessment of tamariki Māori typically assess non-Māori constructs. For example, Goodman's Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997) assesses both conduct issues and prosocial behaviour from a Western worldview. While the SDQ has been validated in Aotearoa across age, gender, ethnicity and deprivation groups (D'Souza et al., 2017; Horwood et al., 2011; Pannekoek & D'Souza, 2018), Māori-specific measures are needed because of concerns about the cultural relevance of Western measures of Māori children (D'Souza et al., 2017). For example, a qualitative study into the cross-cultural acceptability and utility of SDQ reported concerns from Māori parents about the lack of consideration of tamariki Māori in their cultural context and the need for multiple perspectives when interpreting scores (Kersten et al., 2016). A subsequent study evaluating the concurrent validity of the SDQ in comparison to child referral for intervention found that the SDQ had unacceptably low sensitivity in Māori preschool children due to high rates of false positives and, therefore, young Māori children with need for referral were potentially not receiving the appropriate support needed when SDQ was the only method of assessment (Kersten, Vandal, Elder, Tauroa, & McPherson, 2017). Moreover, the 2013 Incredible Years Evaluation report involving young children (Sturrock & Gray, 2013) highlighted concerns about the appropriateness of child and whānau interventions that were not grounded in a Māori worldview nor delivered by Māori and for Māori, an issue that is well documented in the wider literature on Māori identity, well-being and development (Berryman, Macfarlane, & Cavanagh, 2009; Durie, 2004, 2006; McClintock, Mellsop, & Kingi, 2011; McClintock, Tauroa, Mellsop, & Frampton, 2016; Pihama, 2012; Rameka, 2011; G. H. Smith, 2003b).

Given questions about the cultural appropriateness of current measurement tools for tamariki Māori, researchers have argued that the assessment of young Māori children should be culturally relevant, culturally specific and culturally valid, and that measures should be developed by Māori for Māori and reflect Māori realities (Elder, Czuba, Kersten, Caracuel, & McPherson, 2017; Rameka, 2011; Sibley & Houkamau, 2013). It is important for psychology in Aotearoa to understand how best to develop reliable measures to use with tamariki Māori, and how to take into account cultural priorities that may have been overlooked due to presumptions or unconscious bias in mainstream approaches, spanning a range of disciplines (Blank et al., 2016; Pihama, Smith, Taki, & Lee, 2004; G. H. Smith, 2003a; L. T. Smith, 2001). This raises questions about how Indigenous children are unconsciously perceived or stereotyped, thus affecting understandings and judgements during assessment, which in turn might undermine their cultural validity (Blank et al., 2016). Therefore, the development of measurement tools that are not only reliable and valid but also culturally relevant is essential to accurate assessment. These understandings can help foster better clinical practice, as well as informing equitable approaches when working with Māori children and their whānau.

The current study

The current feasibility study, He Piki Raukura, is part of Te Kura Mai i Tawhiti (TKMT), a long-term Māori community-initiated research programme that began in 2012 and has been previously described in detail (Ratima et al., 2019). In brief, the aim of TKMT is to examine the impact that kaupapa Māori early life and whānau programmes have on health, well-being and educational outcomes of the whole whānau. TKMT is a collaboration between Te Pou Tiringa Incorporated and the University of Otago's National Centre for Lifecourse Research. Te Pou Tiringa is the governance entity of Te Kōpae Piripono, a Taranaki Māori-medium early childhood and whānau programme that has been operating since 1994 as an early childhood education centre (ECE). Te Kōpae Piripono was recognised nationally as a 'Centre of Innovation' in 2005 and its programme has previously been described in detail (Tamati, Hond-Flavell, & Korewha, 2008). Te Kōpae Piripono provides a 'real world' kaupapa Māori child and whānau intervention to support and reinforce positive behaviours among young children. The term 'He Piki Raukura' refers to the flight feathers of the toroa (giant albatross), a cherished emblem of the historic Taranaki community of Parihaka as a symbol of peace and of resistance in the face of adversity, and sustained well-being. These are concepts that underpin the work of Te Kōpae Piripono and inform the Māori constructs (Tamati et al., 2021a).

Epistemologically and methodologically the TKMT research programme has a lifecourse orientation and applies an interface approach. This means that the research is located at the interface between the mātauranga Māori and Western science paradigms (Edwards, 2010; Ratima et al., 2019). An interface approach acknowledges that both Māori and Western knowledge systems are equally credible and relevant to the disciplined inquiry in contemporary Aotearoa (Durie, 2004; Edwards, 2003).

The aim of He Piki Raukura has been to both develop and investigate ways to measure Māori constructs underpinning important behaviours in early childhood. In the first phase of the study, interviews were held with whānau and Māori education experts. The Māori researchers then ran a series of wānanga to develop Māori developmental constructs (Tamati et al., 2021a). The four

strengths-based Māori child behaviour constructs identified are – tuakiri (a secure local Māori identity); whānauranga (feeling and acting, as a member of a whānau/community); manawaroa (having courage in adversity, persisting despite difficulty and a positive outlook); and piripono (having integrity, commitment and responsibility for a shared kaupapa/purpose) (Tamati et al., 2021a). In this second feasibility phase of He Piki Raukura, we sought to create a novel set of child behaviour measurement tools based on the above Māori constructs and to determine if these measures were reliable and valid. While other studies have developed Māori measures of identity and wellbeing, following our review of the literature, we concluded that none adequately captured all the necessary elements of the constructs we wished to measure in early childhood. In this stage of the overall study, our aims were to:

1. Develop a set of measurement tools to quantitatively measure identified Māori child behavioural constructs.
2. Test the psychometric properties of the novel measurement tools (i.e., inter-rater reliability, internal consistency and concurrent validity).
3. Refine the measurements tool by developing shortened versions that retain appropriate psychometric properties.

A further aim of He Piki Raukura was to use the validated measures in a third stage, to investigate whether we could detect changes in children's behaviour over the course of 10 months by mapping the trajectories of change in these constructs. This work could only be conducted once the psychometric validation had occurred. The results of this third stage are described in the companion paper, He Piki Raukura - Assessing Ao Māori developmental constructs Part II: Mapping positive change over 10 months among preschool Māori children (Tamati et al., 2021c).

METHODS

Participants

A cohort of 28 children and their 22 immediate whānau (i.e., parents/caregivers) who were enrolled at Te Kōpae Piripono during 2016 agreed to take part in this study. Each whānau was asked to complete a quantitative questionnaire about their child/children and their family at five timepoints, over the course of the 2016 school year (March, June, August, October, December). Data collection occurred across one working week for each of the timepoints. Parents were also asked to consent to their children being videoed over a number of structured activities and also as a part of the day-to-day activities at Te Kōpae Piripono for a rating process described below. In terms of completeness of data, all 28 tamariki and their whānau participated for the entirety of the study, with occasional random missing data due to issues such as illness and tangihanga.

Parents ranged in age from the early twenties to mid-forties (median = 35). However, 88% of parents were aged 27 years and older. Children ranged in age from 11 months to 5 years (median = 3 years 5 months). The gender of the child participants was relatively even (13 boys/15 girls). Mothers made up 81% of adult participants who filled in questionnaires. Twenty-five children attended Te Kōpae Piripono on a full-time basis (35 hours per week). The three children who attended for fewer hours (approximately 30 hours per week) were either younger in age (between 11 months and 15 months) or lived a substantial distance from Te Kōpae Piripono (up to 90 kilometres round trip).

The nominated parent of each child completed the questionnaire at a time and place convenient to them, with one of a team of three research assistants asking the questions (see detailed description of Māori child behaviour questionnaires below). This often happened in families' homes and during weekends. The remainder of whānau completed the questionnaire at Te Kōpae Piripono. Kaitiaki (teachers) at Te Kōpae Piripono also participated in the study. Kaitiaki were randomly allocated a small group of children (approximately N=4) to answer questions about at each of the five timepoints, during data collection. There was a change in one of the kaitiaki at T3 and T4 meaning two other kaitiaki took over rating the children allocated to the original kaitiaki, for T4 and T5. The video observations of children (see detailed description below) were also carried out at Te Kōpae Piripono.

Input and oversight were provided by an expert project advisory group throughout the course of the study. The University of Otago Human Ethics Committee approved the study (16/003). Participants gave informed consent to participate. All of the researchers involved in fieldwork were or had been part of Te Kōpae Piripono in some way, and so there was high trust and strong whānau participation.

Measures

Parents and kaitiaki were asked a series of questions about the children that would best describe their behaviour in relation to the four Māori constructs of interest – tuakiri, whānauranga, manawaroa and piripono. Parents were also asked a series of general demographic questions (e.g., age, gender). A draft questionnaire was composed during the series of wānanga involving the Māori researchers and expert project advisory group and piloted for appropriateness over a seven-month period with relevant whānau in the wider community who were not currently enrolled at Te Kōpae Piripono. During the pilot work, feedback was gathered about the questionnaire's usability and comprehensibility. The resulting Māori Child Behaviour Questionnaire – whānau version (MCBQ-W) and kaitiaki version (MCBQ-K) – measured the four Māori constructs of interest, which are described in detail in Tamati et al. (2021a). For each construct, we generated a set of items that reflected key aspects of that construct. Parents and kaitiaki indicated on a 5-point frequency scale, the extent to which each item in the questionnaire reflected the level of their child's behaviours for each construct (1 = 'not at all'; 2 = 'rarely'; 3 = 'sometimes'; 4 = 'often'; 5 = 'very often'). Parents were asked to rate the four constructs for their children, in three different contexts (i) the home environment (ii) at Te Kōpae Piripono and (iii) in the wider community. Kaitiaki answered questions only in relation to the Te Kōpae Piripono context. All items are available on request.

Parents were also asked to provide feedback on the questionnaire at each timepoint including the extent to

which the questionnaire was easy or difficult to complete, clear or confusing, and appropriate or inappropriate.

Development of the short-form measures

Following data collection, item-total analysis was carried out on the full set of questionnaire items to determine whether it was possible to shorten the multi-item scales. This was to ensure whānau and kaitiaki had clear comprehension of the questions, when rating a child's behaviour, and how they represented a given construct. An item-total correlation test was carried out for both the whānau and kaitiaki ratings and 34 items of the total 199 items (17%) were found to have weak to moderate relationships (r = 0.3-0.4) with the totals of other items. A further 26 items (13%) were removed due to repetition, for being unclear in how they represented a construct, or for not being relevant to a specific context (e.g., one question referred to 'playing in a group environment at home' which was not the case for a number of whānau). Table 1 lists the number of items per construct in the original long-form and the refined short-form. Following this process, the short-form version only was used for the remainder of the analyses.

Child behaviour observations

A series of video observations of children's behaviour were recorded at each timepoint to further evaluate the four constructs of interest and allow testing of convergent validity with the ratings of parents and kaitiaki. This involved videoing children interacting with their peers and kaitiaki during two structured and two unstructured activities (described below) in two different contexts – the kopa kai (dining room) and the kopa mahi (main classroom). Two video cameras, each able to record for a full day, were placed in fixed positions in the kopa kai and the kopa mahi above the whāriki (mat area where most whole group activities occurred (e.g., group reading and kapa haka).

Structured activity #1: Introduction of a new toy

Children were assigned to five groups of approximately five children. The makeup of these groups remained constant for the duration of data collection. Children were randomly selected across mixed age-bands. Each day during data collection week, in the kopa mahi, one of the five groups was introduced to a new toy. Over the course of each data collection phase, all children participated in the activity, at least once. Kaitiaki were asked not to get involved in the play, other than if a child asked for or needed help. The activity lasted for 20 minutes; however, if a child or children spontaneously negotiated for the continued use of the toy then another five minutes was added to the playing time. The introduction of a new toy task sought to elicit children's democratic turn-taking.

Structured activity #2:Pōwhiri (formal welcome)

On two separate days, at each data collection timepoint, a manuhiri (visitor) was welcomed into Te Kōpae Piripono. All children participated in the Taranaki pōwhiri process including harirū first (hongi/shaking of hands) then mihi (words of welcome) and waiata (song), and kai (sharing of food). The video observations, from fixed positions, captured the behaviour and actions of all participating in the welcome process. The pōwhiri provided opportunities to observe children's understanding, behaviour and engagement in tikanga Māori (Māori cultural norms) – including taking on roles, participating in kōrero (speaking) and waiata, assisting others and being able to sit calmly for extended periods.

Unstructured activity #1: Kopa kai (dining room)

The unstructured activities were guided by time sampling principles. The activities in the kopa kai sought to capture children's behaviour during normal meal time activity. Children were randomly selected across age-bands, into three larger groups of between 7-10 children. At morning wā huihui (mat time), each group was assigned a colour e.g. red, green, or yellow group for each of the three dining tables. Each group then ate at the same colour-designated table for the day – across three meal times – kai ata (morning tea), kai poutū (lunch time) and kai ahiahi (afternoon tea). The 'red' table was the table designated to be video recorded. Over three days of the week, each of the groups received a red-coloured card, meaning each group got to sit at the red table at least once.

Unstructured activity #2: Kopa mahi (classroom)

The video camera installed in the kopa mahi was essentially a 'fly on the wall', capturing routine activity during the whole Kōpae day between 9am-3.15pm.

Video rating

The observational data was rated by three researchers, trained to criterion, to rate the Māori constructs of interest. A Māori Child Behaviour Rating Schedule (MCBRS), developed by the research team, was used by the raters (full schedule available on request). The rating given for each of the four constructs was the average rating given across the four different contexts listed above (e.g., structured and unstructured activities). This provided a single rating for each construct (e.g., tuakiri) for each child.

Table 1. Long-form and short-form questionnaire number of items

	Tuakiri		Whanauranga		Manawaroa		Piripono		Total	
	Long	**Short**	Long	**Short**	Long	**Short**	Long	**Short**	Long Total	**Short Total**
Home	10	**8**	16	**10**	13	**9**	10	**7**	49	**34**
Kōpae	10	**7**	16	**11**	13	**9**	10	**7**	49	**34**
Community	13	**10**	16	**11**	13	**9**	10	**7**	52	**37**
Total Whānau	33	**25**	48	**32**	39	**27**	30	**21**	150	**105**
Total Kaitiaki	10	**7**	16	**11**	13	**9**	10	**7**	49	**34**

The MCBRS included a detailed description and characteristics of each construct, as well as observable examples of how a child may display behaviour related to the construct. The schedule also provided instruction on how to rate the observed behaviour on a scale from 1 to 5. A rating of five (5) was given to a child who 'consistently and unprompted, demonstrated examples' of the construct. A rating of four (4) was for a child who 'often, both spontaneously, and sometimes with encouragement by others' exhibited the construct. A rating of three (3) was if a child 'showed some examples of (the construct) with regular encouragement by others.'
A rating of two (2) was if a child demonstrated examples of the construct, 'only if they were reminded or prompted by others and required support to do so'. And a rating of one (1) was if a child demonstrated none of the listed examples of the construct.

Data Analysis

A series of psychometric analyses were conducted to assess inter-rater reliability, internal consistency and concurrent validity of the new measurement tools - the MCBQ-W (whānau questionnaire), MCBQ-K (kaitiaki questionnaire), and the MCBRS (child behaviour rating schedule).

Intra-class correlation coefficients (single measures, one way) were used to measure inter-rater reliability of the video observations. Inter-rater reliability, measured across the three raters at baseline, evaluated how closely aligned their rating were for the same observed child behaviour from the video observations (Cicchetti & Sparrow, 1981; Fleiss, 1981). This process assessed the preliminary ratings and also informed the ongoing training of raters.

Video observations were rated at baseline (T1) for the N=25 children who were enrolled at Te Kōpae Piripono at the time. These ratings were averaged across the four behavioural scenarios (structured activities #1 and #2, and unstructured activities – kopa kai and kopa mahi). A further three children enrolled at T2 and started participating in observational tasks from then on.

Cronbach's alpha were calculated to determine the internal consistency of each measure. This was carried out to show whether the items on each subscale produced similar scores to measure the same underlying constructs (Nunnally & Bernstein, 1994). Cronbach's alphas were calculated individually for the three ratings (child behaviour observations, whānau and kaitiaki ratings) for each of the four Māori constructs over five data collection points (Table 2). Subscales of the MCBQ-W included all three contexts – home, Te Kōpae Piripono and the wider community. A minimum recommended level of alpha coefficients is .70 for preliminary research, .80 for basic research tools and .90 for applied or real-life research – with the ideal being .95 (Nunnally & Bernstein, 1994). Correlations of measures of the same construct were examined to determine associations within multiple measures of the same construct. Factor analyses could not be run with this cohort given the limited number of participants in comparison to the number of items. A regression analysis was carried out for each of the four Māori constructs between the child behaviour ratings (the dependent variable) and the kaitiaki ratings and whānau ratings.

RESULTS

Inter-rater reliability

There was either good or excellent inter-rater reliability for the four constructs, with the intra-class correlation for tuakiri being 0.72 (95% *CI* = 0.46, 0.87); whānauranga 0.65 (95% *CI* = 0.35, 0.83); manawaroa 0.78 (95% *CI* = 0.57, 0.90); and piripono 0.79 (95% *CI* = 0.57, 0.90) (Cicchetti & Sparrow, 1981).

Internal consistency

All four Māori constructs showed very strong internal consistency, with alpha coefficients ranging from .90-.98 (long-form) and .90-.98 (short-form). The scores for the kaitiaki ratings also showed good to excellent internal reliability with alpha coefficients ranging from .89-.97 (long-form) and .88-.97 (short-form). The ratings for the child behaviour video observations showed strong to excellent internal consistency across the four videoed scenarios for all of the four Māori constructs with alpha coefficients ranging from .83-.96.

Within source correlations

There were strong correlations between the scores for the four constructs (Table 3). The constructs were considered to be conceptually distinct based on the kaupapa Māori construct development process, which is described in detail in Tamati et al., (2021a), and involved a series of expert wānanga following qualitative consultation with whānau. We have therefore treated the constructs as separate variables in our analyses despite their inter-correlated scores. However, we still would have expected some relationship between the four constructs, which also aligns with an holistic Māori worldview.

Concurrent validity

Bivariate correlation analyses were carried out between the child behaviour observational ratings and the kaitiaki and whānau ratings, for each of the four Māori constructs. Across the five data collection points, there were generally significant positive associations between the kaitiaki ratings and the child behaviour observations, particularly at T1 and T2 (Table 4). The relationship be tween the majority of the whānau scores and the child behaviour ratings were weak to moderate. Additionally, weakened patterns were experienced for T3 and also for parts of T4 (see Table 4). The reliabilities within subscales remained consistently strong but there was a noticeable dip in the correlations of both the kaitiaki and whānau ratings with the child behaviour observations at T3. There was an increasingly stronger relationship at T4, and at T5 where significant associations for all four constructs were again evident.

Concurrent validity: Regression analyses

The general pattern of the regression analyses (Table 5), indicated that the kaitiaki ratings were likely to be significantly associated to the child behaviour observations, above and beyond the whānau ratings. Again, a dip in associations at T3 was evident in the results. We conducted further correlations and regressions, removing the three children who enrolled at T2, to check whether the same patterns existed for the cohort enrolled for the whole year. The correlations for T3

Table 2. Range of Cronbach's alphas for child behaviour, whānau and kaitiaki ratings

	Tuakiri	Whānauranga	Manawaroa	Piripono
Child observation ratings	.88 - .94	.83 - .93	.88 - .96	.94 - .94
Whānau ratings	.90 - .98	.95 - .97	.94 - .97	.94 - .97
Kaitiaki ratings	.91 - .97	.94 - .97	.92 - .94	.88 - .95

Table 3. Correlations (r) for four Māori child behaviour constructs (sources of information - whānau, kaitiaki, and child observations), at five different timepoints, and the range of magnitude and associated p values, accounting for age.

	Tuakiri	Whānauranga	Manawaroa	Piripono
Whānau				
Tuakiri	1			
Whānauranga	.72, .75, .72, .81, .74 ***p*'s<0.001**	1		
Manawaroa	.58, .56, .32, .49, .48 ***p*'s<0.001-0.101**	.86, .87, .68, .82, .85 ***p*'s<0.001**	1	
Piripono	.59, .74, .77, .80 ,.80 ***p*'s<0.001-0.002**	.82, .88, .88, .93, .93 ***p*'s<0.001**	.67, .82, .67, .78, .77 ***p*'s<0.001**	1
Kaitiaki				
Tuakiri	1			
Whānauranga	.86, .88, .84, .90, .90 ***p*'s<0.001**	1		
Manawaroa	.73, .69, .50, .72, .81 ***p*'s <0.001-.014**	.87, .80, .75, .81, .90 ***p*'s <0.001**	1	
Piripono	.59, .73, .69, .79, .86 ***p*'s <0.001-0.003**	.72, .78, .75, .84, .85 ***p*'s <0.001**	.71, .75, .77, .77, .85 ***p*'s <0.001**	1
Child observations				
Tuakiri	1			
Whānauranga	.91, .94, .88, .94, .92 ***p*'s <0.001**	1		
Manawaroa	.83, .89, .87, .91, .85 ***p*'s <0.001**	.88, .91, .98, .93 ,.91 ***p*'s <0.001**	1	
Piripono	.86, .85, .83, .84, .90 ***p*'s <0.001**	.86, .82, .89, .88, .89 ***p*'s <0.001**	.86, .90, .90, .92, .92 ***p*'s <0.001**	1

Table 4. Correlations of the child behaviour observations with the whānau or kaitiaki ratings

		Correlation with Child Behaviour Observations (r)				
Construct	Source	T1	T2	T3	T4	T5
Tuakiri	Whānau	.644**	.761**	.050	.336	.327
	Kaitiaki	.826**	.840**	-.089	-.083	.464*
Whānauranga	Whānau	.611**	.687**	.086	.445*	.404
	Kaitiaki	.803**	.742**	-.063	.104	.585**
Manawaroa	Whānau	.450*	.564**	.012	.506*	.428*
	Kaitiaki	.755**	.696**	-.038	.031	.635**
Piripono	Whānau	.438*	.588**	-.003	.417	.415
	Kaitiaki	.676**	.739**	-.077	.132	.640**

Table 5. Regression coefficients for associations of child behaviour observations with both the whānau and kaitiaki ratings

	Child Behaviour Observations (beta)				
	T1	T2	T3	T4	T5
Tuakiri					
Whānau ratings	.276	.255	.334	.283	.083
Kaitiaki ratings	.676**	.637**	-.356	.219	.415
Whānauranga					
Whānau ratings	.361**	.404**	.211	.293	.090
Kaitiaki ratings	.668**	.524**	-.196	.257	.532**
Manawaroa					
Whānau ratings	.246	.399**	.033	.481**	.076
Kaitiaki ratings	.682**	.583**	-.052	.068	.590**
Piripono					
Whānau ratings	.336**	.285	.122	.387	-.121
Kaitiaki ratings	.621**	.594**	-.167	.460**	.729**

a. Dependent Variable: Child Behaviour ratings
*** Correlation is significant at the 0.01 level (2-tailed)*
** Correlation is significant at the 0.05 level (2-tailed)*

and T4 looked more similar to the other timepoints (albeit slightly weaker).The regressions showed an association between kaitiaki ratings and the child behaviour observations at T3 but the whānau ratings were more associated with the child behaviour observations at T4, which was in keeping with the whole cohort. These secondary findings are available on request.

DISCUSSION

The behaviour and development of tamariki Māori in Aotearoa and other Indigenous children, globally, are often assessed using purportedly universal child assessment tools created by non-Indigenous researchers, which often decontextualise a child's behaviour (Achenbach & Ruffle, 2000; Corrigan, 2002; Goodman, 1997; Reedtz et al., 2008). Moreover, child assessment has historically taken a deficit-based approach such as identifying conduct problems (Achenbach & Ruffle, 2000; Eyberg & Ross, 1978). Yet research has found that strengths-based assessment approaches are preferred by Māori families (Kersten et al., 2016). An over-reliance on non-Indigenous measurement tools and conceptual approaches to evaluate tamariki Māori risks them being inappropriately evaluated, potentially resulting in them missing out on opportunities for intervention or support that they should be able to access (D'Souza et al., 2017; Kersten et al., 2016). Having appropriate reliable and valid Māori measurement tools is therefore critical in not only reflecting Māori children's cultural backgrounds, but also in providing rich and accurate information about Māori children's development. Such information is a crucial component in the evaluation of kaupapa Māori early years immersion initiatives, which are increasingly recognised as culturally-appropriate and efficacious interventions in Aotearoa.

The current feasibility study, He Piki Raukura, sought to address the lack of Māori measurement tools by taking the four Māori child behaviour constructs of interest – tuakiri, whānauranga, manawaroa and piripono – that had been previously elucidated (Tamati et al., 2021a), and testing them in a cohort of young Māori children attending a kaupapa Māori immersion early years setting. In this, the first of our pair of papers on this overall study, we have described the development of these novel strengths-based Māori child behaviour measurement tools. We then tested the psychometric properties of these measures to determine whether they could reliably assess Māori children's behaviour and also if the measures were meaningful and appropriate to whānau.

We found that the novel measurement tools were internally reliable and concurrently valid. There was strong inter-rater reliability among the video raters. The psychometric properties of the MCBQ-W, MCBQ-K and MCBRS compared favourably with other known measures of young children's behaviour (Corrigan, 2002; Goodman, 2001; Horwood et al., 2011). In our study,

internal consistency for the ratings of the four Māori constructs were strong, which shows that our measurement tools have a similar level of internal consistency to other commonly used tools that provide internally consistent measures of developmental constructs (D'Souza et al., 2017; Ezpeleta, Granero, de la Osa, Penelo, & Domenech, 2013; Gouley, Brotman, Huang, & Shrout, 2008; Horwood et al., 2011; Sturrock & Gray, 2013).

The strong correlations between the scores for the four Māori developmental constructs indicated that the constructs were relatively similar on a statistical level. That is, a child with a high score on one of the constructs was likely to have a high score on the other constructs, particularly whānauranga. This finding could be due to the small number of study participants, suggesting the need for further research using larger cohorts and the use of statistical techniques such as factor analysis. It could also reflect the developmental stage of the children. When we accounted for age, the association between the constructs reduced. Moreover, the Māori constructs are both relational in nature (that is the behaviours were often displayed when children were interacting with each other or with an adult) and conceptually distinct, having been identified through a culturally-grounded construct development process (Tamati et al., 2021a). Also, the child observation tasks intentionally focused on interactions with others. This demonstrates a different worldview approach to that of Western science, which seeks to factor out relationality, rather than embrace it (Kim, Yang, & Hwang, 2006). For example, relatedness is regarded as the 'ultimate premise' of the worldview of Indigenous peoples in Australia (Martin, 2005). Māori researchers, too, argue the importance of relationality, such as whanaungatanga (relationships) and whakapapa (genealogical links with ancestors), atua (Māori deities) and the natural world from a Māori worldview (Bishop, Ladwig, & Berryman, 2014; Macfarlane, Blampied, & Macfarlane, 2011; Rameka, 2011; Wilson-Tukaki & Davis, 2011).

The strong positive correlation between the kaitiaki ratings and the child behaviour observations suggests that the kaitiaki ratings essentially captured a child's behaviour in a similar way as the child behaviour observations. The weaker relationship between some of the whānau scores and the child behaviour ratings shows the whānau ratings provided a slightly different perspective to the child behaviour observations and kaitiaki scores. This is consistent with other findings that show differences in parent and teacher rating of child behaviour (Gao, Paterson, Carter, Iusitini, & Sundborn, 2011; Sargisson, Stanley, & Hayward, 2016), which are often attributed to contextual differences between home and the educational setting, as well as personal and cultural expectations for child behaviour (Gao et al., 2011). In educational and child development literature, teacher and parent views are often sought to examine possible causes or contexts of behaviour, to carry out a whole measurement approach and to explore possible interventions. It is common for parent ratings to be different to that of teachers. Parents see more breadth of their tamariki, across different contexts. Teachers see more of tamariki within the educational setting. Therefore, while different respondents have different insights and perspectives, this does not mean there is no coherence of the factors that are being measured. Rather, it indicates there are different perspectives about a child. This suggests that multiple sources of information provide a more holistic perspective (Gao et al., 2011; Lynne Lane, Stanton-Chapman, Roorbach Jamison, & Phillips, 2007; Sargisson et al., 2016); Sargisson, Stanley, & Hayward, 2016). This diverse information is helpful in fully recognising and building on a child's strengths, skills and abilities, which is a key aim of this research.

For future research, the MCBQ-W and child behaviour observations (MCBRS) were found to be the best combination of measures to use. However, if conducting child behaviour observations is not possible, our findings suggest that the whānau and kaitiaki questionnaires are still reliable to use. A useful process in this feasibility study was the refinement of the original long-form of the questionnaire. This involved removal of some items to reduce repetition and provide greater clarity. For future research, the short-form questionnaire will be quicker to complete, while maintaining the same reliability as the long-form.

We noted a reduction in the correlations between the whānau and kaitiaki ratings and the child behaviour observations at T3 (and somewhat at T4). This may be due to a change of kaitiaki at T3. While the internal reliabilities for all ratings remained consistently strong throughout data collection, the weaker correlations at certain timepoints, indicate the importance of having multiple data collection points (Poulton, Moffitt, & Silva, 2015). In doing so, we were able to better understand potential anomalies, while also identifying relevant factors when conducting research in 'real world' settings.

Based on our review of the literature, we believe this is the first time that child behaviour measurement tools have been created that are grounded in an Indigenous kaupapa Māori worldview. Additionally, these measures have been shown to be psychometrically reliable and valid, meaning they can accurately assess a child in relation to the four constructs of importance to Māori (Tamati et al., 2021a). Therefore, for the first time, researchers in Aotearoa have a reliable set of child behaviour measures from a Māori Indigenous worldview. This means that Māori children can be evaluated or assessed according to their own cultural background.

Importantly, the measures that we created are intentionally strengths-based. The evaluation of Indigenous children has traditionally often been from a deficit-based lens (Dender & Stagnitti, 2011; Fforde, Bamblett, Lovett, Gorringe, & Fogarty, 2013; Rubie-Davies & Peterson, 2016). With a strengths-based approach, we contend that it is still possible to identify children who need help or support, as the rating will show development to the level of a construct. Further, a strengths-based approach aligns with an increasing trend in psychology to move away from deficit approaches to children's development (Craven et al., 2016; Fenton, Walsh, Wong, & Cumming, 2015; Fogarty, Lovell, Langenberg, & Heron, 2018). This not only helps address issues of negative bias toward Māori children (Blank et al., 2016; Pihama et al., 2004), it could also encourage the building of children's strengths and the evaluation of

positive interventions. Having reliable and valid strengths-based, kaupapa Māori measures (Elder et al., 2017; Kersten et al., 2016) of development in young Māori children is crucial, which we have been able to demonstrate with our study.

There are wider positive implications of the potential application of these new measures. The process of assessing young children's development can potentially serve as an evaluation of the quality of their early learning environment and personal contexts to support the development of strengths-based child behaviours. These measures, therefore, can potentially contribute to better ways of evaluating existing kaupapa Māori early years and whānau programmes and interventions (Hond-Flavell, Ratima, Tamati, Korewha, & Edwards, 2017; Ministry of Education, 2013, 2018; Munford, Sanders, Maden, & Maden, 2007; Theodore et al., 2019). These future findings will also help inform government policy and investment, including decisions on when and how prevention and intervention programmes are implemented, as well as for whom (Elder et al., 2017; Harwood et al., 2012; McClintock et al., 2011; Theodore et al., 2019; Treasury New Zealand, 2017).

Strengths and Limitations

The main limitation of this study was its small cohort size (28 children) involving a single Māori early years setting. However, the focus on one cohort of tamariki was intentional in order to carry out the necessarily deep methodological, cultural and practical groundwork, to pilot the measurement tools. The study also required commitment by whānau, kaitiaki and the research staff, as well as generosity of the tamariki. Conducting this type of developmental work across multiple sites, we believe would not have been possible without a high level of trust between all those involved at the centre. Although the number of participants was small, there was sustained whānau involvement throughout the duration of the study. This was assisted by the existing high trust between whānau and the researchers, which it is argued facilitates collaborative inquiry (Cram & Kennedy, 2010).

Strengths of this study include the application of a kaupapa Māori approach to developing Māori child behaviour constructs and measurement tools. These newly created constructs and measurement tools are positioned within a strengths-based framework and they can be used by both whānau and kaitiaki. This is helpful in the context of Aotearoa, as non-deficit assessment approaches are preferred by Māori families (Kersten et al., 2016). Our interface approach to the quantification of Indigenous child development constructs is also a strength of the study. Methodologically, we created child behaviour measurement tools from an Ao Māori perspective. In keeping with our interface approach (Edwards, 2010), we also utilised widely used psychometric processes to test these measures. In this way, the research has drawn from the strengths of mātauranga Māori and Western science knowledge systems to generate new knowledge and about measurement of Māori developmental constructs.

Concluding Comments

We hope that our research process will be a useful model to other groups of kaupapa Māori researchers and Māori communities seeking to build an evidence-base around their own programmes using our measures of the four Māori constructs or to develop measures that tap into constructs of meaning to them. The development of Māori measurement tools like this can enable Māori communities to test psychometrically sound measures and their relationship to positive life outcomes. In an accompanying paper (Tamati et al., 2021c), we examine changes over time based on the data collected, to test whether our measurement tools can detect meaningful change in the four constructs over 10 months, during a school year.

We are mindful that this is a feasibility study, so future work is needed with larger cohorts of tamariki Māori to continue validating our measures. There is exciting potential to trial these measures in other Māori and possibly other Indigenous contexts. We remain hopeful that this research will offer alternative, more authentic and robust approaches to working with Māori children and whānau to improve their life outcomes.

References

Achenbach, T. M., & Ruffle, T. M. (2000). The child behavior checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in Review, 21*(8), 265-271. doi:10.1542/pir.21-8-265

Allen, J., Mohatt, G. V., Fok, C. C. T., Henry, D., Burkett, R., & Team, P. A. (2014). A protective factors model for alcohol abuse and suicide prevention among Alaska Native youth. *American Journal of Community Psychology, 54*(1-2), 125-139. doi:10.1007/s10464-014-9661-3

Batool, S. S., & Khalid, R. (2011). Development of indigenous scale of emotional intelligence and evaluation of its psychometric properties. *Pakistan Journal of Social and Clinical Psychology, 9*, 66. Retrieved from https://www.researchgate.net/profile/Ruhi_Khalid/publication/309558024_Development_of_Indigenous_Scale_of_Emotional_Intelligence_and_Evaluation_of_Its_Psychometric_Properties/links/58171fe708aedc7d8967b66d/Development-of-Indigenous-Scale-of-Emotional-Intelligence-and-Evaluation-of-Its-Psychometric-Properties.pdf

Berryman, M., Macfarlane, S., & Cavanagh, T. (2009). Indigenous contexts for responding to challenging behaviour: Contrasting western accountability with maori restoration of harmony. *International Journal of Restorative Justice, 5*(1), 1. Retrieved from https://search.proquest.com/openview/b8c6c1bb9dd7577 4db689e723be70721/1?pq-origsite=gscholar&cbl=75969

Bishop, R., Berryman, M., Cavanagh, T., & Teddy, L. (2009). Te kotahitanga: Addressing educational disparities facing Māori students in New Zealand. *Teaching and Teacher Education, 25*(5), 734-742. doi:10.1016/j.tate.2009.01.009

Bishop, R., Ladwig, J., & Berryman, M. (2014). The centrality of relationships for pedagogy. *American Educational Research Journal, 51*(1), 184-214. doi:10.3102/0002831213510019

Blank, A., Houkamau, C., & Kingi, H. (2016). *Unconscious bias and education: A comparative study of Māori and African American students*: Oranui Diversity Leadership.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency 86* (2), 127-137.

Corrigan, A. (2002). Social Competence Scale–Parent Version Grade 1/Year 2. In *Fast Track Project Technical Report*: Available from the Fast Track Project Website http://www.fasttrackproject.org.

Cram, F., & Kennedy, V. (2010). Researching with whānau collectives. *Mai Review, 3*, 1-12. Retrieved from https://www.researchgate.net/profile/Fiona_Cram2/publication/49594477_Researching_with_Whanau_Collectives/links/584f2f8308aecb6bd8d027c5/Researching-with-Whanau-Collectives.pdf

Craven, R. G., Ryan, R. M., Mooney, J., Vallerand, R. J., Dillon, A., Blacklock, F., & Magson, N. (2016). Toward a positive psychology of indigenous thriving and reciprocal research partnership model. *Contemporary Educational Psychology, 47*, 32-43. doi:10.1016/j.cedpsych.2016.04.003

D'Souza, S., Waldie, K. E., Peterson, E. R., Underwood, L., & Morton, S. M. (2017). Psychometric properties and normative data for the preschool strengths and difficulties questionnaire in two-year-old children. *Journal of Abnormal Child Psychology, 45*(2), 345-357. doi:10.1007/s10802-016-0176-2

Dender, A., & Stagnitti, K. (2011). Development of the Indigenous Child-Initiated Pretend Play Assessment: Selection of play materials and administration. *Australian Occupational Therapy Journal, 58*(1), 34-42. doi:10.1111/j.1440-1630.2010.00905.x

Durie, M. (2004). *Exploring the interface between science and indigenous knowledge.* Paper presented at the 5th APEC Research and Development Leaders Forum, Christchurch, New Zealand.

Durie, M. (2006). Measuring māori wellbeing. *New Zealand Treasury Guest Lecture Series, 1*.

Durie, M., Cooper, R., Grennell, D., Snively, S., & Tuaine, N. (2010). Whānau ora: Report of the taskforce on whānau-centred initiatives. *To: Hon Tariana Turia Minister for the Community and Voluntary Sector*.

Edwards, W. J. W. (2003). *Te ihu waka: The interface between research and Māori development: A thesis presented in partial fulfillment of the requirements for the degree of Master of Philosophy at Massey university.* Massey University, Retrieved from http://hdl.handle.net/10179/6499

Edwards, W. J. W. (2010). *Taupaenui: Maori positive ageing.* Massey University, Retrieved from http://hdl.handle.net/10179/1331

Elder, H., Czuba, K., Kersten, P., Caracuel, A., & McPherson, K. (2017). Te Waka Kuaka, Rasch analysis of a cultural assessment tool in traumatic brain injury in Māori [version 1; referees: 1 approved with reservations, 1 not approved]. *F1000Research, 6*(1034). doi:10.12688/f1000research.11500.1

Eyberg, S. M., & Ross, A. W. (1978). Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child Psychology, 7*(2), 113-116. doi:10.1080/15374417809532835

Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E., & Domenech, J. M. (2013). Psychometric properties of the strengths and difficulties questionnaire(3-4) in 3-year-old preschoolers. *Compr Psychiatry, 54*(3), 282-291. doi:10.1016/j.comppsych.2012.07.009

Fenton, A., Walsh, K., Wong, S., & Cumming, T. (2015). Using strengths-based approaches in early years practice and research. *International Journal of Early Childhood, 47*(1), 27-52. doi:10.1007/s13158-014-0115-8

Fergusson, D. M., Horwood, J., Ridder, E., & Grant, H. (2005). *Early Start evaluation report*: Early Start Project Limited.

Fforde, C., Bamblett, L., Lovett, R., Gorringe, S., & Fogarty, B. (2013). Discourse, deficit and identity: Aboriginality, the race paradigm and the language of representation in contemporary Australia. *Media International Australia, 149*(1), 162-173. doi:10.1177/1329878X1314900117

Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement, 5*(1), 105-112.

Fogarty, W., Lovell, M., Langenberg, J., & Heron, M.-J. (2018). Deficit discourse and strengths-based approaches: changing the narrative of Aboriginal and Torres Strait Islander health and wellbeing. *Deficit Discourse and Strengths-based Approaches: Changing the Narrative of Aboriginal and Torres Strait Islander Health and Wellbeing*, viii. Retrieved from https://search.informit.com.au/documentSummary;dn=583838279439901;res=IELIND

Gao, W., Paterson, J., Carter, S., Iusitini, L., & Sundborn, G. (2011). Agreement and discordance of parents' and teachers' reports of behavioural problems among Pacific children living in New Zealand. *AUT Pacific Islands Families Study Of those Born in 2000, at Manukau City, New Zealand, 17*(2), 65.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(11), 1337-1345. Retrieved from https://ac.els-cdn.com/S0890856709605438/1-s2.0-S0890856709605438-main.pdf?_tid=3d404711-ab56-44e0-abeb-7e435cadcc29&acdnat=1533260112_8e185985bdd39e95d75f9f5e721f5a1d

Gouley, K. K., Brotman, L. M., Huang, K.-Y., & Shrout, P. E. (2008). Construct Validation of the Social Competence Scale in Preschool-age Children. *Social Development, 17*(2), 380-398. doi:10.1111/j.1467-9507.2007.00430.x

Harwood, M., Weatherall, M., Talemaitoga, A., Alan Barber, P., Gommans, J., Taylor, W., . . McNaughton, H. (2012). An assessment of the Hua Oranga outcome instrument and comparison to other outcome measures in an intervention study with Maori and Pacific people following stroke. *NZ Medical Journal, 125* (1364), 55-67. Retrieved from http://www.nzma.org.nz/journal/read-the-journal/all-issues/2010-2019/2012/vol-125-no-1364/article-harwood

Haswell, M. R., Kavanagh, D., Tsey, K., Reilly, L., Cadet-James, Y., Laliberte, A., . . . Doran, C. (2010). Psychometric validation of the Growth and Empowerment Measure (GEM) applied with Indigenous Australians. *Australian and New Zealand Journal of Psychiatry, 44*(9), 791-799. doi:10.3109/00048674.2010.482919

Hond, R. (2013). *Matua te reo, Matua te tangata: Speaker community: Visions, approaches, outcomes.* (PhD). Massey University, Palmerston North. Retrieved from http://hdl.handle.net/10179/5439

Hond-Flavell, E., Ratima, M., Tamati, A., Korewha, H., & Edwards, W. (2017). Te Kura Mai i Tawhiti: He Tau Kawekaweā: Building the foundation for whanau educational success and wellbeing; a Kaupapa Māori

ECE approach. *Teaching & Learning Research Initiative*. Retrieved from http://www.tlri.org.nz/tlri-research/research-completed/ece-sector/te-kura-mai-i-tawhiti-he-tau-kawekawe%C4%81-building

Horwood, L. J., Gray, D. S., & Fergusson, D. (2011). *The Psychometric Properties of the Child Behaviour Rating Scales used in the Incredible Years Pilot Study*. Unpublished.

Houkamau, C. A., & Sibley, C. G. (2010a). Māori cultural efficacy and subjective wellbeing: A psychological model and research agenda. *Social Indicators Research, 103*(3), 379-398. doi:10.1007/s11205-010-9705-5

Houkamau, C. A., & Sibley, C. G. (2010b). The Multi-dimensional Model of Maori Identity and Cultural Engagement. *New Zealand Journal of Psychology, 39*(1), 8-28. Retrieved from www.psychology.org.nz/journal-archive/NZJP-Vol391-2010-2-Houkamau.pdf

Housman, A., Dameg, K., Kobashigawa, M., & Brown, J. (2011). Report on the Hawaiian oral language assessment (H-OLA) development project. *Second Language Studies, 29*(2), 1-59. Retrieved from www.hawaii.edu/sls/wp-content/uploads/2014/09/Housman-et-al.pdf

Keown, L. J., Sanders, M. R., Franke, N., & Shepherd, M. (2018). Te Whānau Pou Toru: a Randomized Controlled Trial (RCT) of a Culturally Adapted Low-Intensity Variant of the Triple P-Positive Parenting Program for Indigenous Māori Families in New Zealand. *Prevention Science, 19*(7), 954-965. doi:10.1007/s11121-018-0886-5

Kersten, P., Dudley, M., Nayar, S., Elder, H., Robertson, H., Tauroa, R., & McPherson, K. M. (2016). Cross-cultural acceptability and utility of the strengths and difficulties questionnaire: views of families. *BMC Psychiatry, 16*(1), 347. doi:10.1186/s12888-016-1063-7

Kersten, P., Vandal, A. C., Elder, H., Tauroa, R., & McPherson, K. M. (2017). Concurrent Validity of the Strengths and Difficulties Questionnaire in an Indigenous Pre-School Population. *Journal of Child and Family Studies, 26*(8), 2126-2135. doi:10.1007/s10826-017-0725-5

Kim, U., Yang, K.-S., & Hwang, K.-K. (2006). Contributions to indigenous and cultural psychology. In *Indigenous and Cultural Psychology* (pp. 3-25): Springer.

King, A., & Turia, T. (2002). *He korowai oranga*: Ministry of Health.

LoGiudice, D., Strivens, E., Smith, K., Stevenson, M., Atkinson, D., Dwyer, A., . . . Flicker, L. (2011). The KICA Screen: the psychometric properties of a shortened version of the KICA (Kimberley Indigenous Cognitive Assessment). *Australasian Journal of Ageing, 30*(4), 215-219. doi:10.1111/j.1741-6612.2010.00486.x

Lynne Lane, K., Stanton-Chapman, T., Roorbach Jamison, K., & Phillips, A. (2007). Teacher and Parent Expectations of Preschoolers' Behavior: Social Skills Necessary for Success. *Topics in Early Childhood Special Education, 27*(2), 86-97. doi:10.1177/02711214070270020401

Macfarlane, A. H., Blampied, N. M., & Macfarlane, S. H. (2011). Blending the clinical and the cultural: A framework for conducting formal psychological assessment in bicultural settings. *New Zealand Journal of Psychology, 40*(2). Retrieved from https://pdfs.semanticscholar.org/161e/200fba124c980a38fcb618c28663780b4c2e.pdf

Mane, J. (2009). Kaupapa Māori: A community approach. *Mai Review, 3*, 1. Retrieved from http://www.review.mai.ac.nz/mrindex/MR/article/download/243/243-1710-1-PB.pdf

Martin, K. (2005). Childhood, lifehood and relatedness: Aboriginal ways of being, knowing and doing. *Introductory indigenous studies in education: The importance of knowing*, 27-40.

McClintock, K., Mellsop, G. W., & Kingi, T. K. R. (2011). Development of a culturally attuned psychiatric outcome measure for an indigenous population. *International Journal of Culture and Mental Health, 4*(2), 128-143. doi:10.1080/17542863.2010.537484

McClintock, K., Tauroa, R., Mellsop, G., & Frampton, C. (2016). Pilot of Te Tomo mai, a child and adolescent mental health service evaluation tool for an indigenous rangatahi (youth) population. *International Journal of Adolescence and Youth, 21*(1), 96-103. doi:10.1080/02673843.2013.813861

Ministry of Education. (2013). *Ka Hikitia – Accelerating for Success 2013-2017*. Wellington: Ministry of Education

Ministry of Education. (2018). *Draft strategic plan for early learning 2019-29*. Wellington: Ministry of Education. conversation.education.govt.nz

Morton, S., Grant, C., Berry, S. D., Walker, C. G., Corkin, M., Ly, K., . . . Bandara, D. K. (2017). Growing Up in New Zealand: A longitudinal study of New Zealand children and their families. Now We Are Four: Describing the preschool years. Retrieved from http://ro.uow.edu.au/sspapers/3130

Munford, R., Sanders, J., Maden, B., & Maden, E. (2007). Blending whanau/family development, parent support and early childhood education programmes. *Social Policy Journal of New Zealand, 32*, 72-87. Retrieved from ww.minedu.govt.nz/index.cfm?layout=index&indexid=10987&indexparentid=1094

Nunnally, J. C., & Bernstein, I. (1994). Validity. *Psychometric Theory, 3*, 99-132.

Palmer, S. (2004). Homai te Waiora ki Ahau: A tool for the measurement of wellbeing among Maori-the evidence of construct validity.

Pannekoek, L., & D'Souza, S. (2018). *Psychometric properties of the Strengths and Difficulties Questionnaire in a multi-ethnic population of pre- and primary school-aged children*. Unpublished.

Peterson, E. R., Dando, E., D'Souza, S., Waldie, K. E., Carr, A. E., Mohal, J., & Morton, S. M. (2018). Can Infant Temperament Be Used to Predict Which Toddlers Are Likely to Have Increased Emotional and Behavioral Problems? *Early Education and Development, 29*(4), 435-449. doi:10.1080/10409289.2018.1457391

Pihama, L. (2012). Kaupapa Māori theory: Transforming theory in Aotearoa. *He Pukenga Korero, 9*(2). Retrieved from http://www.hepukengakorero.com/index.php/HPK/article/viewFile/2/pdf

Pihama, L., Smith, K., Taki, M., & Lee, J. (2004). A literature review on kaupapa Māori and Māori education pedagogy. *The International Research Institute for Maori and Indigenous Education*. Retrieved from http://hdl.voced.edu.au/10707/212867

Ponitz, C. E. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23*(2), 141-158. doi:10.1016/j.ecresq.2007.01.004

Poulton, R., Moffitt, T. E., & Silva, P. A. (2015). The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the

future. *Social psychiatry and psychiatric epidemiology, 50*(5), 679-693. doi:10.1007/s00127-015-1048-8

Rameka, L. (2011). Being māori: Culturally relevant assessment in early childhood education. *Early Years, 31*(3), 245-256. doi:10.1080/09575146.2011.614222

Ratima, M., Theodore, R., Tamati, A., Hond-Flavell, E., Edwards, W., Korewha, H., . . . Poulton, R. (2019). Te Kura Mai i Tawhiti Research Programme: A collaborative lifecourse approach to health, wellbeing and whānau development. *MAI Journal, 8,* 63-76. doi:10.20507/MAIJournal.2019.8.1.5

Reedtz, C., Bertelsen, B., Lurie, J., Handegård, B. H., Clifford, G., & MØRCH, W. T. (2008). Eyberg Child Behavior Inventory (ECBI): Norwegian norms to identify conduct problems in children. *Scandinavian Journal of Psychology, 49*(1), 31-38. doi:10.1111/j.1467-9450.2007.00621.x

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at three to seven years: The Children's Behavior Questionnaire. *Child Development, 72*(5), 1394-1408. doi:10.1111/1467-8624.00355

Royal Tangaere, A. (2012). *Te hokinga ki te ukaipō: A socio-cultural construction of Māori language development: Kōhanga Reo and home.* ResearchSpace@ Auckland, Retrieved from http://hdl.handle.net/2292/13392

Rubie-Davies, C. M., & Peterson, E. R. (2016). Relations between teachers' achievement, over-and underestimation, and students' beliefs for Māori and Pākehā students. *Contemporary Educational Psychology, 47*, 72-83. doi:doi.org/10.1016/j.cedpsych.2016.01.001

Sargisson, R. J., Stanley, P. G., & Hayward, A. (2016). Multi-informant scores and gender differences on the Strengths and Difficulties Questionnaire for New Zealand children. *New Zealand Journal of Psychology (Online), 45*(2), 4. Retrieved from https://pdfs.semanticscholar.org/33f7/a025cdf3ff80eef9540773abc764f910fba9.pdf

Schlesinger, C. M., Ober, C., McCarthy, M. M., Watson, J. D., & Seinen, A. (2007). The development and validation of the Indigenous Risk Impact Screen (IRIS): a 13-item screening instrument for alcohol and drug and mental health risk. *Drug and Alcohol Review, 26*(2), 109-117. doi:10.1080/09595230601146611

Sibley, C. G., & Houkamau, C. A. (2013). The Multi-Dimensional Model of Maori Identity and Cultural Engagement: Item Response Theory Analysis of Scale Properties. *Cultural Diversity & Ethnic Minority Psychology, 19*(1), 97-110. doi:10.1037/a0031113

Smith, G. H. (2003a). Indigenous struggle for the transformation of education and schooling. *Transforming Institutions: Reclaiming Education and Schooling for Indigenous Peoples*, 1-14. Retrieved from http://www.rangahau.co.nz/assets/Smith,%20G/indigenous_struggle.pdf

Smith, G. H. (2003b). *Transforming institutions : reclaiming education and schooling for Indigenous peoples.* Paper presented at the Alaskan Federation of Natives Convetion, Anchorage, Alaska, USA.

Smith, L. T. (2001). Decolonizing Methodologies, Research and Indigenous Peoples, 3rd impression. In: Zed Books Ltd./University of Otego Press, London, New York, Dunedin.

Sturrock, F., & Gray, D. (2013). Incredible Years pilot study evaluation report. *Wellington: Centre for Research and Evaluation, Ministry of Social Development*.

Sturrock, F., Gray, D., Fergusson, D., Horwood, J., & Smits, C. (2014). Incredible Years Follow-up Study– Long-term follow-up of the New Zealand Incredible Years Pilot Study. *Ministry of Social Development*. Retrieved from http://www.incredibleyears.com/wp-content/uploads/indredible-years-follow-up-study-new-zealand-2014.pdf

Tamati, A., Hond-Flavell, E., & Korewha, H. (2008). Te Kōpae Piripono Centre of Innovation Research Report. In *Education Counts*. Wellington: Ministry of Education.

Tamati, A., Ratima, M., Hond-Flavell, E., Edwards, W., Hond, R., Korewha, H., Theodore, R., Treharne, G.J. & Poulton, R. (2021a). He Piki Raukura: Understanding Strengths-based Māori Child Development Constructs in Kaupapa Māori Early Years Provision. *MAI Journal, 10(1),* 17-29. http://www.journal.mai.ac.nz/sites/default/files/MAI_Jrnl%202021_V10_I1_Tamati_FINAL.pdf

Tamati, A., Treharne, G. J., Kokaua, J., Theodore, R., Ratima, M., Hond-Flavell, E., Edwards, W., Hond, R. & Poulton, R. (2021c). He Piki Raukura: Assessing Ao Māori developmental constructs - Part II: Mapping positive change over 10 months in preschool Māori children. *New Zealand Journal of Psychology, 50*(2), 35-45.

Theodore, R., Ratima, M., Edwards, W., Sporle, A., Te Morenga, L., Kiro, C., & Hond, R. (2019). How a lifecourse approach can promote long-term health and wellbeing outcomes for Māori. *Journal of Indigenous Wellbeing, 4*(1), 15-25. Retrieved from https://journalindigenouswellbeing.com/media/2019/07/122.126.How-a-lifecourse-approach-can-promote-long-term-health-and-wellbeing-outcomes-for-M%C4%81ori.pdf

Treasury New Zealand. (2017). *Social investment*. Retrieved from http://www.treasury.govt.nz/statesector/socialinvestment

Wilson-Tukaki, A., & Davis, K. (2011). Relating to others: Three layers of knowing. *Early Childhood Folio, 15*(2), 20. Retrieved from https://search.informit.com.au/documentSummary;dn=742034742344782;res=IELHSS

Affiliations

Aroaro Tamati:
Taranaki Iwi, Ngāti Ruanui, Te Atiawa, Te Whānau-ā-Apanui

Mihi Ratima:
Ngāti Awa, Whakatōhea

Erana Hond-Flavell:
Taranaki Iwi, Ngāti Ruanui, Te Atiawa, Te Whānau-ā-Apanui

Will Edwards:
Taranaki Iwi, Ngāruahine, Tāngahoe, Pakakohi, Ngāti Ruanui

Ruakere Hond:
Taranaki Iwi, Ngāti Ruanui, Te Atiawa, Te Whānau-ā-Apanui

Reremoana Theodore:
Ngāpuhi

Corresponding Author

Aroaro Tamati
Te Pou Tiringa Incorporated
PO Box 6016, Moturoa
New Plymouth 4344

Aotearoa/New Zealand
Email: aroaro.tamati@gmail.com

Acknowledgements
We would like to sincerely thank the whānau of Te Kōpae Piripono and members of our Oversight Group who include Professor Sue Crengle, Professor Mera Penehira and Professor Leonie Pihama. We acknowledge funding for this research received from the Health Research Council of New Zealand in the form of a Ngā Kanohi Kitea Full Project Grant [Grant 13/954] and a Feasibility Grant [Grant 16/587]; in addition Reremoana Theodore was supported by a Māori Health Emerging Leader Fellowship [Grant 18/644]; Mihi Ratima was supported by a Ngā Pou Senior Fellowship [Grant 16/440]; and Ruakere Hond was supported by a Hōhua Tūtengaehe Research Fellowship [Grant 16/586]. We would also like to acknowledge the following who have provided funding for this research: the Ministry of Education, TOI Foundation (formerly TSB Community Trust), the University of Otago, the New Zealand Council for Educational Research, the National Centre for Lifecourse Research, and Te Pou Tiringa Incorporated.

Māori Glossary

Ao Māori	Māori world; Māori worldview
Aotearoa	Indigenous name for New Zealand
He Piki Raukura	One of the projects of Te Kōpae Piripono's longitudinal research, that focusses on Māori child behavioural constructs
Hapū	sub-tribe
Hariru	handshake
Hongi	Māori cultural greeting
Iwi	tribe
Kai	food
Kai ahiahi	afternoon tea
Kai ata	morning tea
Kaitiaki	teacher at Te Kōpae Piripono
Kapa Haka	Māori cultural form of dance
Kaupapa	purpose, objective, topic, philosophy
Kaupapa Māori	Māori philosophical framework
Kopa kai	dining room
Kopa mahi	classroom
Kōpae	shortened name of Te Kōpae Piripono (Taranaki-based Māori immersion early childhood centre)
Kōrero	speak; speaking
Manawaroa	the notion of having courage in adversity, persisting despite difficulty and a positive outlook
Mātauranga Māori	Māori Indigenous knowledge systems
Mihi	greeting; speech of acknowledgement
Parihaka	historic Māori settlement south of New Plymouth, NZ
Piripono	the notion of having integrity, commitment and responsibility for a shared kaupapa/purpose
Pōwhiri	ceremonial Māori welcome
Tamariki	children
Tamariki Māori	Māori children
Taranaki	a region in the west of the North Island; a tribe
Te Kōpae Piripono	Taranaki-based Māori immersion early years and whānau initiative
Te Kura Mai i Tawhiti	the name given to Te Kōpae Piripono's longitudinal research programme
Te Pou Tiringa	governing board of Te Kōpae Piripono
Te reo Māori	Māori language
Toroa	giant albatross
Tikanga Māori	Māori process, customs,
Tuakiri	the notion of a secure local Māori identity
Waiata	song; singing
Wānanga	Māori cultural process of knowledge generation and learning
Whakapapa	genealogy; genealogical connection
Whānau	family, usually encompassing wider membership than the nuclear family
Whāriki	mat
Whānauranga	the notion of feeling and acting, as a member of a whānau/community

He Piki Raukura: Assessing Ao Māori developmental constructs – Part II: Mapping positive change over 10 months among preschool Māori children

Aroaro Tamati[1,2], Gareth J. Treharne[2], Jesse Kokaua[2], Reremoana Theodore[2], Mihi Ratima[1,2], Erana Hond-Flavell[1,2], Will Edwards[1,2], Ruakere Hond[1,2], Richie Poulton[2]

[1] Te Pou Tiringa Incorporated, New Plymouth, Taranaki, Aotearoa/New Zealand

[2] National Centre for Lifecourse Research, Department of Psychology, University of Otago, Aotearoa/New Zealand

Supporting positive early childhood development is important for both short and long-term outcomes. This paper is the second of two papers examining the measurement of four Māori constructs underpinning positive child behaviours – tuakiri (secure local Māori identity); whānauranga (acting as a member of a whānau); manawaroa (persisting despite difficulty); and piripono (having integrity, commitment and responsibility). Here, we describe changes in measures of these constructs over a 10 month period. Whānau (families) and kaitiaki (teachers) completed questionnaires and video observations were made of 28 Māori children aged 0-5 years. Growth curve analysis revealed significant positive change in each construct across five timepoints, even controlling for age differences. These findings provide proof-of-concept that our novel measures of the four constructs are sensitive to change in positive child behaviours among preschool Māori children.

Keywords: *Māori child behaviour, Māori child development, kaupapa Māori, early childhood assessment, indigenous measurement tools, growth curve models*

Introduction

Evidence generated over recent decades demonstrates positive effects of well-designed preschool education programmes on children's developmental and behavioural outcomes in the short-term, and on their potential life trajectories over the longer term (Bakken, Brown, & Downing, 2017; Barnett, 2008; Camilli, Vargas, Ryan, & Barnett, 2010; Isaacs & Roessel, 2008; Watts, Gandhi, Ibrahim, Masucci, & Raver, 2018; Wylie & Thompson, 2003). Early childhood is regarded as a critical period where the foundations of children's emotional well-being and social development are laid (Bakken et al., 2017; Barnett, 2008; F. A. Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; K. Campbell, Chen, Shenoy, & Cunningham, 2019; Gomajee et al., 2018; McCoy et al., 2017; Schindler et al., 2015). In Aotearoa, early years interventions are deemed pivotal to positive life outcomes (Horwood, Gray, & Fergusson, 2011; Sturrock, Gray, Fergusson, Horwood, & Smits, 2014). Furthermore, kaupapa Māori early years initiatives, grounded within Māori worldviews, are considered appropriate interventions that can improve life outcomes for Māori (Durie, Cooper, Grennell, Snively, & Tuaine, 2010; King & Turia, 2002). To date, there are few published evaluations of such programmes. Given the wide and enduring inequities in education for Māori (Bishop, Berryman, Cavanagh, & Teddy, 2009; Bishop & Glynn, 2003; Mahuika, Berryman, & Bishop, 2011) there is an urgent need for studies investigating the impacts of early years kaupapa Māori approaches. Our study seeks to help address this gap.

Recent international research about the positive long-term benefits of early years programmes, is mixed. Past research has identified several methodological features that confound interpretation including: a) reductions in early academic advantage once children attend school (Ansari, 2018; Claessens & Garrett, 2014; Hill, Gormley Jr, & Adelstein, 2015; Schimmel, 2018); b) older studies that do not account for current educational policies and practices; and c) evaluations that have to navigate differences in early years settings, learning standards, curriculums, programme quality and access for young children (Black et al., 2017; Cortázar, 2015; Felfe & Lalive, 2018). Another limitation of the research to date is that many studies have tended to focus on socially and economically disadvantaged children and address deficits or delinquent behaviours in children, rather than positive behaviours or strengths-based approaches (Bakken et al., 2017; Burger, 2010; Murray et al., 2018; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). Furthermore, it has been argued that previous research evaluating early learning approaches has lacked thorough analysis of diverse curriculum content and culturally responsive teaching practices (Schimmel, 2018). These issues highlight the need for more systematic evaluation of strengths-based culturally-centred indigenous early years programmes.

Longitudinal studies where participants are examined using the same measurement tools at regular timepoints over a designated period can be used to identify changes in child development (McArdle & Nesselroade, 2003; Zhang, McArdle, & Nesselroade, 2012). One way of doing this is the technique of growth curve modelling to explore the impact of interventions, as well as describing developmental trajectories. There are a number of studies that have used growth curve modelling to evaluate change (or 'growth') in children's development over a series of timepoints. Most of this research has tended to focus on longitudinal academic outcomes, such as numeracy and literacy (Burchinal, Peisner-Feinberg, Pianta, & Howes, 2002; Hojnoski, Silberglitt, & Floyd, 2009). Growth curve modelling has also been used to track conduct problems and internalising behaviours (Mathiesen, Sanson, Stoolmiller, & Karevold, 2009; Yew & O'Kearney, 2015), as well as adaptive learning behaviours and prosocial trajectories (Domínguez, Vitiello, Maier, & Greenfield, 2010; Gomajee et al., 2018). Interestingly, some studies of preschool children have not found gender differences (Knafo & Plomin, 2006; Mathiesen et al., 2009; Sharp, Croudace, Goodyer, & Amtmann, 2005; Venker, Ray-Subramanian, Bolt, & Weismer, 2014). Age, on the other hand, has been found to be a factor in significant positive change over time for a range of constructs in young children, including social competence, autism, hyperactivity/inattention and conduct problems (Domínguez et al., 2010; Gomajee et al., 2018; Santos, Vaughn, Peceguina, Daniel, & Shin, 2014; Venker et al., 2014).

Māori children are typically tested and evaluated using Western-derived measurement tools, such as the social competence scale (Horwood et al., 2011), Parents' Evaluations of Developmental Status (Glascoe, 1997), and the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). A systematic review of the psychometric properties of the SDQ raised questions about its cultural validity including parent disquiet about its deficit approach (Kersten et al., 2016). To attain cultural validity in measures for tamariki Māori, there is a need to include measures that are developed by Māori, for Māori (Elder & Kersten, 2015). Issues have also been raised about the SDQ's concurrent validity in preschool Māori children, highlighting concerns about developmental or behavioural problems going undetected and children not getting the support they need (Kersten, Vandal, Elder, Tauroa, & McPherson, 2017). The above issues raise questions about the types of measures and approaches currently used to screen young Māori children and why Māori-created measurement tools that focus on behaviours of interest for Māori have yet to be developed and applied.

One local research project that has used the SDQ is the longitudinal study, Growing Up in New Zealand (GUINZ), where ethnic disparities and mother-reported behavioural difficulties were identified and examined (D'Souza, Underwood, Peterson, Morton, & Waldie, 2019; Morton et al., 2018). The GUINZ study was set up to be broadly representative of the national demographic by including 'ethnically diverse' participants from four groups – European, Māori, Pasifika and Asian (Morton et al., 2017; Peterson et al., 2018, p. 436). In an exploration of persistence and change, D'Souza et al. (2019) found that most young children who displayed 'abnormal' behaviours at age two typically improved by 4.5 years, while a significant proportion still persisted in these difficulties, over the same period. Also reported was an association between ethnicity and behavioural difficulties, particularly for Māori or Pasifika children (D'Souza et al., 2019; Peterson et al., 2018). European ethnicity was a consistent predictor of fewer behavioural problems, such as hyperactivity, emotion, peer and conduct problems, among the 2 year olds (Peterson et al., 2018, p. 444). According to D'Souza et al. (2019) this could have been due to greater exposure to certain factors such as early adversity or socioeconomic disadvantage, prompting the authors to advise that the results be treated with caution. They also argued the need for further research to explore the underlying 'causes' of such ethnic disparity including broader social and historical factors.

In the current feasibility study, we address the gap that remains in child behaviour research in Aotearoa by intentionally working within a Māori worldview (Rameka, 2011, 2012; Sibley & Houkamau, 2013). Māori researchers have long argued the importance of culturally valid pedagogies, measures and assessments that 'make sense' to Māori (Bishop, 1999; Mahuika et al., 2011; Rameka, 2011, p. 246; Sibley & Houkamau, 2013). There are also limitations of research intended to make a difference in children's lives that does not include cultural and ethnic groups (Denham et al., 2003). The design of assessment tools and measures for young Māori children needs to mirror 'culturally-located interpretive systems, and that these are different for Māori and non-Māori,' (Rameka, 2012, p. 34). Based on our searches of the literature we determined that there have been no longitudinal studies specifically on the behaviour of tamariki Māori, that have used a strengths-based approach or have focussed on Māori constructs of children's behaviour.

Our previous work has identified four Māori child behaviour constructs – tuakiri (a secure local Māori identity); whānauranga (feeling and acting, as a member of a whānau/community); manawaroa (having courage in adversity, persisting despite difficulty and a positive outlook); and piripono (having integrity, commitment and responsibility for a shared kaupapa/purpose) (Tamati et al., 2021a). This paper documents a continuation of the study, He Piki Raukura, which is part of a longitudinal Māori community-initiated research programme, Te Kura Mai i Tawhiti (TKMT) (for further details see Ratima et al., 2019). A key aim of the overall research programme is to explore the impact of kaupapa Māori early life and whānau programmes on positive life outcomes for young children and their whānau. An important step in achieving this long-term objective is to test whether culturally-relevant outcome measures of the four Māori child behaviour constructs are sensitive to change over time. Having created, tested and validated our measurement tools in the previous paper in this series (Tamati et al., 2021b) we next sought to identify whether we could measure change in behaviour over a period of 10 months.

preschool tamariki Māori would show an increase over the five timepoints across all four constructs (i.e., intra-individual change with a child acting as their own

Table 1: Participation over five timepoints

T1 7-11 Mar			T2 13-17 Jun			T3 29Aug-2Sep			T4 31 Oct-4 Nov			T5 5-9 Dec		
W	K	**R**	W	K	R	W	K	R	W	K	R	W	K	R
25	25	25	28	26	23	28	28	23	26	23	22	28	23	22

W= Māori Child Behaviour Whānau Questionnaire, K = Māori Child Behaviour Kaitiaki Questionnaire, R = Māori Child Behaviour Rating Schedule

control). We also hypothesised that there would be differences in the changes over time in the four Māori constructs by age and gender. We do not have directional hypotheses for these associations as this is the first study of these constructs.

METHODS

Participants

As described in the accompanying paper, He Piki Raukura: Assessing Ao Māori developmental constructs - Part I: Reliability of novel strengths-based measures among preschool Māori children (Tamati et al., 2021b), this study involved the same 28 tamariki, aged between 11 months and 5 years of age, and their immediate whānau. These children were enrolled at Te Kōpae Piripono, a Taranaki-based Māori-medium early childhood programme (Tamati, Hond-Flavell, & Korewha, 2008), during the 2016 school year. Parents were aged from 22 years to 41 years (median 35). The median age of the children was 3 years 5 months. Fourteen girls and 12 boys participated in the research. Most of the children were enrolled full-time (i.e., 35 hours per week). A small number of children attended for fewer hours (approximately 30 hours per week) due to being younger or because of distance (i.e., some children travelled up to 90 kilometres each day to attend).

The study received ethics approval from the University of Otago Human Ethics Committee - 16/003. All participating whānau gave written consent. Whānau members answered questions about their tamariki and their family at five timepoints in the year: T1 (7-11 March), T2 (13-17 June), T3 (29 August - 2 September), T4 (31 October - 4 November) and T5 (5-9 December). One parent from each whānau agreed to answer the questions at each of the data points (81% of these were the mothers). Kaitiaki (teachers) of Te Kōpae Piripono answered questions about tamariki who were randomly allocated to them. During these data collection periods, tamariki were also video recorded participating in both structured and day-to-day activities at Te Kōpae Piripono. All 28 tamariki and their whānau enrolled at Te Kōpae Piripono participated in the study. At T1, 25 children and their whānau participate, and three additional children from one whānau started at T2. Full data for all measures was collected on at least 19 of the tamariki at any one of the five timepoints. Occasional data points are missing at random and were not imputed.

Measures

We applied a repeated measures design, whereby our novel measurement tools were used at each of the five data collection timepoints. These measures included the Māori Child Behaviour Questionnaire – whānau version (MCBQ-W) and kaitiaki (teacher) version (MCBQ-K) – and the Māori Child Behaviour Rating Schedule (MCBRS), which are described in detail in the accompanying paper (Tamati et al., 2021b). Our previous analyses demonstrated the reliability and concurrent validity of the measures (Tamati et al., 2021b). Parents and kaitiaki rated items that described children's behaviour in relation to the four Māori constructs of interest – tuakiri, whānauranga, manawaroa and piripono (Tamati et al., 2021b). Additionally, the video observations were recorded and rated at a later date on the same four constructs. The video ratings were carried out by three researchers, trained to criterion on the Māori constructs of interest, using the MCBRS. An average rating was given for each of these constructs across different contexts including structured and unstructured activities. An analysis of the consistency across video raters found good to excellent inter-rater reliability for each of the four constructs (Tamati et al., 2021b).

Data Analysis

Data analysis involved individually modelling the growth in each of the four constructs over time to test the significance of any changes over the five timepoints. A sequence of linear mixed effects growth models was conducted, as described by Ghisletta, Renaud, Jacot, and Courvoisier (2015). This included a random effects component for variation across individuals, namely for age and gender. Age was analysed as three levels, according to children's year of birth, that is, for children born in 2010/11, 2012/13, and 2014/15. The aim of this approach was to represent children's developmental phases.

A step-wise approach was used to first test a direct relationship, in terms of whether the outcome variables (tuakiri, whānauranga, manawaroa and piripono) increased over time. Next, we tested whether age or gender of the children influenced the way they responded over the five measurement phases. The model sequence was: Unadjusted Model 1, a mixed effects growth model across different phases with independent residuals across phases; Adjusted Model 2, as for Model 1 with age added; and Adjusted Model 3, as for model 2 with gender also added.

RESULTS

The unadjusted and adjusted rates of change over time increased significantly over the five timepoints for all four Māori constructs across all three sources of information – the whānau ratings on the MCBQ-W, kaitiaki ratings on the MCBQ-K and child behaviour observational ratings on the MCBRS - with the single exception of kaitiaki ratings for tuakiri, which did not increase significantly (p=0.175) (Table 2).

After controlling for age or age and gender, the patterns of significance changed little in the four constructs over the five timepoints for whānau, kaitiaki and child behaviour observational ratings (Table 2). The adjustment for children's age led to a slight reduction in the coefficients for increase in mean scores across timepoints (adjusted 2). Further adjustment for gender made little change to the coefficients for increase in mean scores across timepoints (Table 2, adjusted 3).

Associations for age and gender reported in the models adjusting for these two variables (adjusted 3) were generally significant for age across the four Māori constructs but less often for gender differences (Table 3). The association of age with whānau rating of the constructs was positive, as indicated by significantly negative coefficients (because age was coded in terms of year of birth), meaning the older the child the higher the average scores. The direction of age influence was not consistent for the other sources of ratings. Older children had higher whānauranga and piripono scores across all sources of ratings.

While few significant gender differences were observed overall, kaitiaki ratings and child behaviour ratings were significantly lower for girls than boys for tuakiri and manawaroa constructs. This is in contrast to other gender differences that, while not significant, were all in the opposite direction.

DISCUSSION

Kaupapa Māori early years initiatives are argued as having considerable potential to make a difference in the lives of tamariki Māori (Durie et al., 2010; King & Turia, 2002). However, there are few studies evaluating children's development in kaupapa Māori early years settings, particularly studies that follow children over time. Such evaluation matters. Wide and enduring inequities persist for Māori across wide-ranging areas including wealth, health, wellbeing and education (Bishop et al., 2009; Mahuika et al., 2011; Ratima & Jenkins, 2012). Our feasibility study contributes preliminary evidence that can help address this gap in knowledge. We sought to establish proof-of-principle that four newly developed and validated measures of key Māori constructs – tuakiri, whānauranga, manawaroa and piripono – were able to detect positive change through time among children attending a Māori immersion early child educational setting (Te Kōpae Piripono). Our findings showed the anticipated developmental progression for all four Māori constructs, even when controlling for age and gender. This suggests that increases in positive behaviour might result from other factors including the content, approach and character of kaupapa Māori early years settings, that foster the development of these constructs from a very young age and regardless of gender.

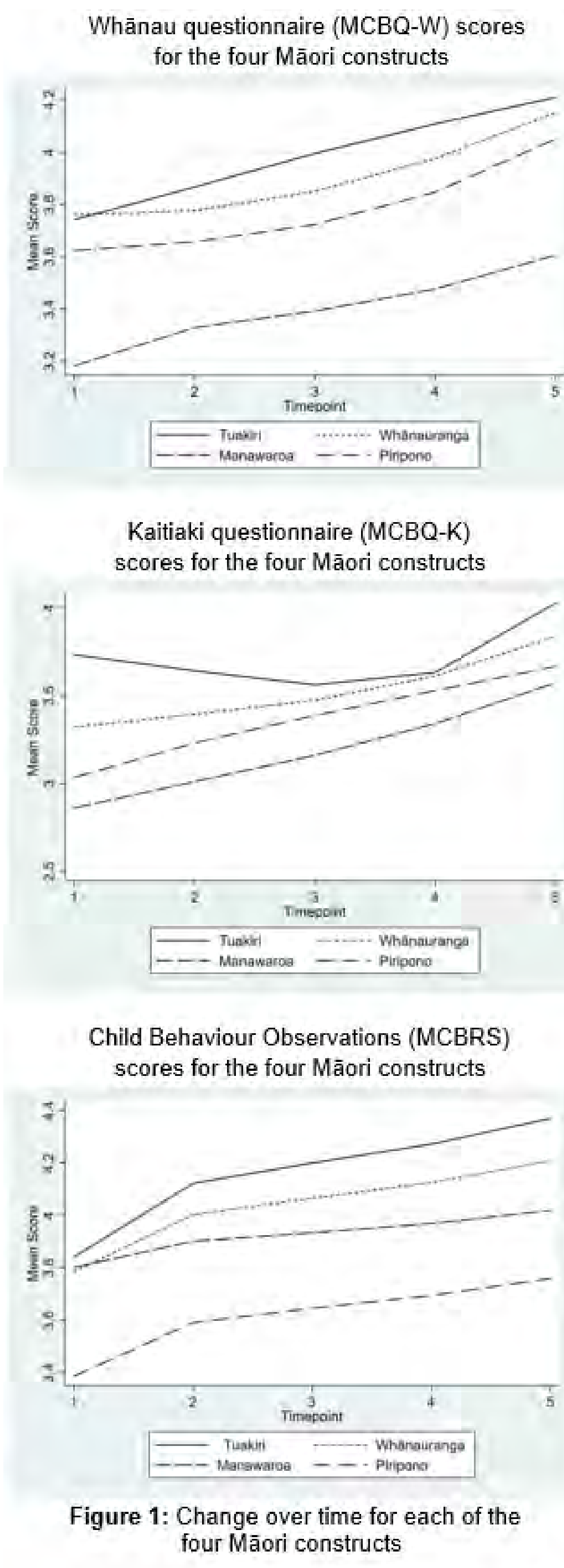


Figure 1: Change over time for each of the four Māori constructs

Table 2: Model coefficients and p-values for change over time for whānau, kaitiaki, and child behaviour observational ratings (unadjusted and adjusted for age and gender).

		Whānau (MCBQ-W)			Kaitiaki (MCBQ-K)			Child Behaviour Observational Ratings (MCBRS)		
		Coef.	***z***	***P*>|*z*|**	**Coef.**	***z***	***P*>|*z*|**	**Coef.**	***z***	**P>|*z*|**
Tuakiri	Unadjusted	0.0834	2.68	0.007	0.0465	1.36	0.175	0.1425	4.08	<0.0001
	Adjusted 2 *for age*	0.0811	2.69	0.007	0.0599	1.7	0.089	0.1547	4.73	<0.0001
	Adjusted 3 *age & gender*	0.0816	2.7	0.007	0.0609	1.73	0.083	0.1549	4.76	<0.0001
Whānauranga	Unadjusted	0.0821	5.19	<0.0001	0.1270	5.45	<0.0001	0.0000	0	<0.0001
	Adjusted 2 *for age*	0.0797	5.25	<0.0001	0.1354	6.11	<0.0001	0.1404	4.04	<0.0001
	Adjusted 3 *age & gender*	0.0801	5.26	<0.0001	0.1400	6.24	<0.0001	0.1417	4.08	<0.0001
Manawaroa	Unadjusted	0.0803	4.18	<0.0001	0.1893	7.93	<0.0001	0.0905	2.53	0.012
	Adjusted 2 *for age*	0.0790	4.13	<0.0001	0.1957	7.86	<0.0001	0.1042	3.19	0.001
	Adjusted 3 *age & gender*	0.0791	4.13	<0.0001	0.1976	7.79	<0.0001	0.1033	2.91	0.004
Piripono	Unadjusted	0.0752	3.14	0.002	0.1441	5.21	<0.0001	0.1192	3.39	0.001
	Adjusted 2 *for age*	0.0728	3.06	0.002	0.1478	5.21	<0.0001	0.1387	4.09	<0.0001
	Adjusted 3 *age & gender*	0.0728	3.06	0.002	0.1508	5.36	<0.0001	0.1394	4.14	<0.0001

Table 3: The coefficient of mean scores, associated with age and gender

		Whānau			Kaitiaki			Child Behaviour Observational Ratings		
		Coef.	*z*	*P*>|*z*|	Coef.	*z*	*P*>|*z*|	Coef.	*z*	*P*>|*z*|
Tuakiri	Age	-0.2389	-3.25	0.001	0.0609	1.73	0.083	0.1549	4.76	<0.0001
	Gender	0.1148	0.53	0.595	-0.5256	-4.49	<0.0001	-0.4168	-5.56	<0.0001
Whānauranga	Age	-0.2580	-4.40	<0.0001	-0.3744	-4.81	<0.0001	-0.4075	-5.33	<0.0001
	Gender	0.1446	0.84	0.398	0.4152	1.93	0.053	0.1702	0.78	0.436
Manawaroa	Age	-0.2289	-3.31	0.001	0.1976	7.79	<0.0001	0.1033	2.91	0.004
	Gender	0.0418	0.21	0.837	-0.4054	-6.00	<0.0001	-0.4611	-6.35	<0.0001
Piripono	Age	-0.3460	-4.82	<0.0001	-0.4811	-7.16	<0.0001	-0.5686	-6.73	<0.0001
	Gender	0.0074	0.04	0.972	0.2367	1.23	0.217	0.2391	0.99	0.323

The association with age in the current study is similar to other studies where, as to be expected, age has been found to be a factor in significant change in young children for a range of behavioural and neurodevelopmental constructs such as social competence, autism, hyperactivity/inattention and conduct problems (Domínguez et al., 2010; Gomajee et al., 2018; Santos et al., 2014; Taylor, Eisenberg, Spinrad, Eggum, & Sulik, 2013; Venker et al., 2014). Most of these studies are deficit-based and focus predominantly on problem behaviours. In contrast, our study is strengths-based and emphasises prosocial behaviours from an Ao Māori worldview, in the form of the four Māori constructs.

Our findings about gender differences for the four Māori constructs are preliminary but suggest there are few significant differences between boys and girls and only on the ratings by kaitiaki or of the video observation and not the whānau ratings. This general lack of gender differences is consistent with findings from a number of studies (Hojnoski et al., 2009; Knafo & Plomin, 2006; Mathiesen et al., 2009; Sharp et al., 2005; Venker et al., 2014). One longitudinal study using growth curve modelling that is similar to ours reported strong evidence for the importance of genetics and environment in the development of prosocial behaviour in 9424 sets of twins, aged from 2 to 7 of age and that except for what they describe as the 'usual finding' of a mean difference favouring girls in terms of prosocial behaviours, the results for gender were 'strikingly similar' (Knafo & Plomin, 2006, p. 782).

The increase in the scores on each of the four Māori constructs over time is an important finding. This means that we are able to see growth in the levels of strengths-based child behavioural constructs - tuakiri, whānauranga, manawaroa and piripono - that are strongly grounded in te ao Māori. This is first such demonstration as a far as we are aware. Moreover, we had hypothesised directional change (i.e. improvements represented by higher scores) across the 10-month period and this was confirmed. Critically, the change appeared to occur independent of normal ageing/developmental processes. That is, some developmental milestones appear to occur as function of maturation or aging (Lightfoot, Cole, & Cole, 2018). This could not be assumed for these specific constructs, hence the novelty and utility of this finding. With regard to the latter, the findings suggest that these constructs can now be safely usefully incorporated into larger interventions in which appropriate comparison groups are included.

These constructs take a clearly different approach to behaviour, with a particular focus on collectivity and relationality, rather than individuality (Brayboy, Gough, Leonard, Roehl, & Solyom, 2012; Higgins & Kim, 2019; Smith & Smith, 2018; Wilson & Kamana, 2001). Take the construct of whānauranga, for example. In terms of child development, this means that a child is displaying observable behaviours that demonstrate they increasingly feel and act as a member of a whānau or community. Observable behaviours for whānauranga include prosocial characteristics such as manaaki and tiaki - showing respect, generosity, sharing and caring, and being helpful, reassuring and kind to others (Hond, 2017; Moorfield, 2019; Noddings, 2012; H. W. Williams, 1957). These characteristics resonate with other Western prosocial constructs including self-regulation, mindfulness and kindness (Flook, Goldberg, Pinger, & Davidson, 2015) self-control (Moffitt et al., 2011); social-emotional skills (Nix, Bierman, Domitrovich, & Gill, 2013) decision making and self-efficacy (Domitrovich, Durlak, Staley, & Weissberg, 2017); ego-resiliency and empathy (Taylor et al., 2013) and patience (Barragan-Jason & Atance, 2017). The above research indicates that these types of behaviour have positive impacts for children.

Strengths and Limitations

The main limitation of this overall study was that it comprised a single sample, thereby constraining the generalisability of the findings beyond this specific early childhood setting (Te Kōpae Piripono). However, we thought it prudent to provide 'proof-of-principle' i.e. sensitivity to change through time, before enrolling participants in a larger and more ambitious study in which comparisons are made across different types of service provision. Our group was not large, however the total enrolled tamariki at Te Kōpae Piripono (N=28) participated and there were high levels of commitment over the 10 month period. It was also noteworthy that there was sustained whānau participation for the duration of the study, providing a comprehensive picture of all the tamariki and their whānau who took part.

The main goal in this study was to test the feasibility of using Maori constructs to capture change over time, in order to carry out future work on a larger scale. Although the sample size may appear low, power calculations (available on request) revealed that we had sufficient numbers to detect significant change over time, while controlling for age. This shows that, even with small numbers of children, medium to large effects could be detected for all the Māori constructs. This outcome is reflected in similarly-sized studies (of children aged 3-5 years) where significant growth was also detected (K. Campbell et al., 2019; Kan & Kohnert, 2012).

The comparatively short duration of the study (i.e. data collected over 10 months) is also a potential limitation of this work because it represents a relatively brief snapshot of development. On the other hand, showing that it was possible to measure change in this reasonably short period speaks to the sensitivity and potential utility of the scales. Additionally, this feasibility study provided a much needed focus on 'deeper' measurement by applying a multi-measure, multi-source assessment approach, which is an established process for longitudinal research (Poulton, Moffitt, & Silva, 2015).

We acknowledge that, from a Western research methods perspective, the non-randomised methodology we applied in this study may have limitations in terms of how definitive the conclusions can be. However, our interface research approach accesses two legitimate knowledge systems. Randomised-control methodologies raise complicated questions about tikanga Māori including the concept of whai wāhi, an imperative of inclusion of all participants (Moorfield, 2019).

There are a also number of strengths of the study. First, the study involved the exploration of a unique set of constructs that were created within a Māori/Indigenous worldview (Tamati et al., 2021a). Second, the work was also innovative in taking an interface approach and

applying psychometric methods to explore measurement of the novel Māori constructs (Durie, 2004; Edwards, 2010; Nakata, 2016; Ratima et al., 2019). This interface approach enables the research to draw on the strengths of both mātauranga Māori and Western science approaches to address research questions in a way that leverages from both knowledge systems as being of relevance in kaupapa research Māori. Third, as noted in the accompanying paper on the psychometrics of our novel measures (Tamati et al., 2021b), our study has an intentionally strengths-based orientation. This is in contrast to considerable literature that has tended to focus on identifying or analysing problem behaviours (Achenbach & Ruffle, 2000; Goodman, 1997; Honig, 2009; Keenan & Wakschlag, 2000; L. R. Williams et al., 2009). However, our study is in keeping with current trends in psychology where there is more of an emphasis on children's abilities and also on positive approaches to children's development (Aspinwall & Staudinger, 2003; Craven et al., 2016; Fenton, Walsh, Wong, & Cumming, 2015).

Implications for policy and practice

This study's value is based on use of reliable and culturally-appropriate Māori child behaviour measurement tools to begin to explore ways kaupapa Māori early years approaches may impact on the development of positive Māori child behaviours. In doing so, the study breaks new ground by providing evaluation and assessment tools to other researchers looking to work with Māori children and their whānau. This is in line with the long-held call from Māori researchers that Māori-focussed childhood research is needed (Neha, 2016). The tools may also potentially be used to help generate an evidence base around the importance and value of quality kaupapa Māori early years initiatives.

Our research demonstrates innovative ways of examining how indigenous constructs that can be practically operationalised and measured in whānau and Māori community settings, particularly using quantitative methods and in relation to children's behaviour. Furthermore, as well as developing novel tools to measure the short-term impact of quality (Meade, 2010; Ritchie, 2008) early years kaupapa Māori programmes on children's development, the present feasibility study lays the groundwork for further research. In particular, there is potential for exploring the longer-term effects of the development and expression of children's behaviours and the positive impact of kaupapa Māori interventions over time (Hond-Flavell, Ratima, Tamati, Korewha, & Edwards, 2017; Ministry of Education, 2013, 2018; Munford, Sanders, Maden, & Maden, 2007; Theodore et al., 2019).

Future research is needed to expand on this feasibility study by systematically assessing tamariki changes on these important constructs in a larger comparison study. This future work will ultimately provide a richer evidence base about programmes or approaches that are effective for Māori, resulting in the long-term positive outcomes which are an essential goal for research in this field (Ratima et al., 2019).

This study is also important in its contribution to the development of novel Māori theoretical frameworks for lifecourse research, as well as Māori-specific research processes, protocols and tools. The work has relevance for early years kaupapa Māori early childhood education provision in Aotearoa, and for indigenous peoples and communities internationally. Indeed, there is potential that this study can help create the best evidence for early years interventions for all New Zealanders, as per the ongoing goals of Vision Mātauranga and of Māori researchers (Ministry of Research Science and Technology, 2007; Rauika Māngai, 2020). Specifically, this study derives new knowledge, distinctly grounded in an Ao Māori worldview, that advances mātauranga about culturally-appropriate ways of working with and assessing young Māori children, acknowledging and respecting children's own cultural backgrounds and supporting children's development from their own cultural lens. This is important, particularly as approximately 80 percent of Māori children attend mainstream early years education settings (Ministry of Education, 2019).

References

Achenbach, T. M., & Ruffle, T. M. (2000). The child behavior checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in Review, 21*(8), 265-271. doi:10.1542/pir.21-8-265

Ansari, A. (2018). The persistence of preschool effects from early childhood through adolescence. *Journal of Educational Psychology, 110*(7), 952. doi:10.1037/edu0000255

Aspinwall, L. G., & Staudinger, U. M. (2003). A psychology of human strengths: Some central issues of an emerging field. In L. G. Aspinwall & U. M. Staudinger (Eds.), *A Psychology of Human Strengths: Fundamental Questions and Future Directions for a Positive Psychology* (pp. 9-22). Washington, DC: American Psychological Association c2003.

Bakken, L., Brown, N., & Downing, B. (2017). Early Childhood Education: The Long-Term Benefits. *Journal of Research in Childhood Education, 31*(2), 255-269. doi:10.1080/02568543.2016.1273285

Barnett, W. S. (2008). *Preschool education and its lasting effects: Research and policy implications*. Retrieved from Center & Education Policy Research Unit: Arizona State Unviersity, University of Colorado: http://epicpolicy.org/publication/preschool-education

Barragan-Jason, G., & Atance, C. M. (2017). A new approach to measuring patience in preschoolers. *Developmental Psychobiology, 59*(6), 738-748. doi:10.1002/dev.21535

Bishop, R. (1999). *Kaupapa maori research: An indigenous approach to creating knowledge*. Paper presented at the Maori psychology: Research and practice, University of Waikato.

Bishop, R., Berryman, M., Cavanagh, T., & Teddy, L. (2009). Te kotahitanga: Addressing educational disparities facing Māori students in New Zealand. *Teaching and Teacher Education, 25*(5), 734-742. doi:10.1016/j.tate.2009.01.009

Bishop, R., & Glynn, T. (2003). *Culture counts: Changing power relations in education*: Zed Books.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., . . . Shiffman, J. (2017). Early childhood development coming of age: science through the life course. *The Lancet, 389*(10064), 77-90. doi:10.1016/S0140-6736(16)31389-7

Brayboy, B. M., Gough, H. R., Leonard, B., Roehl, R. F., & Solyom, J. A. (2012). Reclaiming scholarship: Critical Indigenous research methodologies. *Qualitative Research: An Introduction to Methods and Designs*, 423-450.

Burchinal, M. R., Peisner-Feinberg, E., Pianta, R., & Howes, C. (2002). Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of School Psychology, 40*(5), 415-436. doi:10.1016/S0022-4405(02)00107-3

Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly, 25*(2), 140-165. doi:10.1016/j.ecresq.2009.11.001

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record, 112*(3), 579-620.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science, 6*(1), 42-57. doi:10.1207/S1532480XADS0601_05

Campbell, K., Chen, Y.-J., Shenoy, S., & Cunningham, A. E. (2019). Preschool children's early writing: Repeated measures reveal growing but variable trajectories. *Reading and Writing, 32*(4), 939-961. doi:10.1007/s11145-018-9893-y

Claessens, A., & Garrett, R. (2014). The role of early childhood settings for 4–5 year old children in early academic skills and later achievement in Australia. *Early Childhood Research Quarterly, 29*(4), 550-561. doi:10.1016/j.ecresq.2014.06.006

Cortázar, A. (2015). Long-term effects of public early childhood education on academic achievement in Chile. *Early Childhood Research Quarterly, 32*, 13-22. doi:10.1016/j.ecresq.2015.01.003

Craven, R. G., Ryan, R. M., Mooney, J., Vallerand, R. J., Dillon, A., Blacklock, F., & Magson, N. (2016). Toward a positive psychology of indigenous thriving and reciprocal research partnership model. *Contemporary Educational Psychology, 47*, 32-43. doi:10.1016/j.cedpsych.2016.04.003

D'Souza, S., Underwood, L., Peterson, E. R., Morton, S. M., & Waldie, K. E. (2019). Persistence and change in behavioural problems during early childhood. *BMC Pediatrics, 19*(1), 259. doi:10.1186/s12887-019-1631-3

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach–Major, S., & Queenan, P. (2003). Preschool emotional competence: Pathway to social competence? *Child Development, 74*(1), 238-256.

Domínguez, X., Vitiello, V. E., Maier, M. F., & Greenfield, D. B. (2010). A longitudinal examination of young children's learning behavior: Child-level and classroom-level predictors of change throughout the preschool year. *School Psychology Review, 39*(1), 29-47. doi:10.1080/02796015.2010.12087788

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development, 88*(2), 408-416.

Durie, M. (2004). *Exploring the interface between science and indigenous knowledge.* Paper presented at the 5th APEC Research and Development Leaders Forum, Christchurch, New Zealand.

Durie, M., Cooper, R., Grennell, D., Snively, S., & Tuaine, N. (2010). Whānau ora: Report of the taskforce on whānau-centred initiatives. *To: Hon Tariana Turia Minister for the Community and Voluntary Sector*.

Edwards, W. J. W. (2010). *Taupaenui: Maori positive ageing.* Massey University, Retrieved from http://hdl.handle.net/10179/1331

Elder, H., & Kersten, P. (2015). Whakawhiti kōrero, a method for the development of a cultural assessment tool, Te Waka Kuaka, in Māori traumatic brain injury. *Behavioural Neurology, 2015*. doi:10.1155/2015/137402

Felfe, C., & Lalive, R. (2018). Does early child care affect children's development? *Journal of Public Economics, 159*, 33-53. doi:10.1016/j.jpubeco.2018.01.014

Fenton, A., Walsh, K., Wong, S., & Cumming, T. (2015). Using strengths-based approaches in early years practice and research. *International Journal of Early Childhood, 47*(1), 27-52. doi:10.1007/s13158-014-0115-8

Flook, L., Goldberg, S. B., Pinger, L., & Davidson, R. J. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Developmental Psychology, 51*(1), 44.

Ghisletta, P., Renaud, O., Jacot, N., & Courvoisier, D. (2015). Linear mixed-effects and latent curve models for longitudinal life course analyses. In C. Burton-Jeangros, S. Cullati, A. Sacker, & D. Blane (Eds.), *A life course perspective on health trajectories and transitions* (pp. 155-178): Springer, Cham.

Glascoe, F. P. (1997). Parents' evaluation of developmental status (PEDS). *Ellsworth & Vandermeer Press, Ltd*.

Gomajee, R., El-Khoury, F., Côté, S., Van Der Waerden, J., Pryor, L., & Melchior, M. (2018). Early childcare type predicts children's emotional and behavioural trajectories into middle childhood. Data from the EDEN mother–child cohort study. *Journal of Epidemiology and Community Health, 72*(11), 1033-1043.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Higgins, M., & Kim, E.-J. (2019). De/colonizing methodologies in science education: rebraiding research theory–practice–ethics with Indigenous theories and theorists. *Cultural Studies of Science Education, 14*(1), 111-127. doi:10.1007/s11422-018-9862-4

Hill, C. J., Gormley Jr, W. T., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? *Early Childhood Research Quarterly, 32*, 60-79. doi:10.1016/j.ecresq.2014.12.005

Hojnoski, R. L., Silberglitt, B., & Floyd, R. G. (2009). Sensitivity to growth over time of the preschool numeracy indicators with a sample of preschoolers in Head Start. *School Psychology Review, 38*(3), 402-418. doi:10.1080/02796015.2009.12087823

Hond, R. (2017). Tikanga Tuku Iho - Parihaka Legacy Statement. *Parihaka/Crown Reconciliation Ceremony Programme*(Unpublished), Parihaka Papakāinga Trust. Parihaka:Taranaki.

Hond-Flavell, E., Ratima, M., Tamati, A., Korewha, H., & Edwards, W. (2017). Te Kura Mai i Tawhiti: He Tau Kawekaweā: Building the foundation for whanau educational success and wellbeing; a Kaupapa Māori ECE approach. *Teaching & Learning Research Initiative*. Retrieved from http://www.tlri.org.nz/tlri-

research/research-completed/ece-sector/te-kura-mai-i-tawhiti-he-tau-kawekawe%C4%81-building

Honig, A. S. (2009). Understanding and working with non-compliant and aggressive young children. *Early Child Development and Care, 179*(8), 1007-1023. doi:10.1080/03004430701726217

Horwood, L. J., Gray, D. S., & Fergusson, D. (2011). *The Psychometric Properties of the Child Behaviour Rating Scales used in the Incredible Years Pilot Study*. Unpublished.

Isaacs, J. B., & Roessel, E. (2008). Impacts of early childhood programs. *Washington: First Focus and Brookings Institution*.

Kan, P. F., & Kohnert, K. (2012). A growth curve analysis of novel word learning by sequential bilingual preschool children. *Bilingualism: Language and Cognition, 15*(3), 452-469.

Keenan, K., & Wakschlag, L. S. (2000). More than the terrible twos: The nature and severity of behavior problems in clinic-referred preschool children. *Journal of Abnormal Child Psychology, 28*(1), 33-46. doi:10.1023/A:1005118000977

Kersten, P., Dudley, M., Nayar, S., Elder, H., Robertson, H., Tauroa, R., & McPherson, K. M. (2016). Cross-cultural acceptability and utility of the strengths and difficulties questionnaire: views of families. *BMC Psychiatry, 16*(1), 347. doi:10.1186/s12888-016-1063-7

Kersten, P., Vandal, A. C., Elder, H., Tauroa, R., & McPherson, K. M. (2017). Concurrent Validity of the Strengths and Difficulties Questionnaire in an Indigenous Pre-School Population. *Journal of Child and Family Studies, 26*(8), 2126-2135.

King, A., & Turia, T. (2002). *He korowai oranga*: Ministry of Health.

Knafo, A., & Plomin, R. (2006). Prosocial behavior from early to middle childhood: genetic and environmental influences on stability and change. *Developmental Psychology, 42*(5), 771. doi:10.1037/0012-1649.42.5.771

Lightfoot, C., Cole, M., & Cole, S. R. (2018). *The development of children*: Worth Publishers.

Mahuika, R., Berryman, M., & Bishop, R. (2011). Issues of culture and assessment in New Zealand education pertaining to Maori students. *Assessment Matters, 3*, 183. Retrieved from https://search.informit.com.au/documentSummary;dn=516054067390993;res=IELNZC>

Mathiesen, K. S., Sanson, A., Stoolmiller, M., & Karevold, E. (2009). The Nature and Predictors of Undercontrolled and Internalizing Problem Trajectories Across Early Childhood. *Journal of Abnormal Child Psychology, 37*(2), 209-222. doi:10.1007/s10802-008-9268-y

McArdle, J. J., & Nesselroade, J. R. (2003). Growth Curve Analysis in Contemporary Psychological Research. In *Handbook of Psychology* (pp. 447-480).

McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., . . . Shonkoff, J. P. (2017). Impacts of Early Childhood Education on Medium- and Long-Term Educational Outcomes. *Educational Researcher, 46*(8), 474-487. doi:10.3102/0013189x17737739

Meade, A. (2010). *The contribution of ECE Centres of Innovation to building knowledge about teaching and learning 2003–2010.* Paper presented at the Paper presentedto TLRI EarlyYears Symposium. Wellington, NZ: Teaching and Learning ResearchInitiative. Nuttall, J.(2010, November). Thecontribution of The Teaching And.

Ministry of Education. (2013). *Ka Hikitia – Accelerating for Success 2013-2017*. Wellington: Ministry of Education

Ministry of Education. (2018). *Draft strategic plan for early learning 2019-29*. Wellington: Ministry of Education. conversation.education.govt.nz

Ministry of Education. (2019). Attendance of licensed ECE services in 2019. *Early Childhood Education Census 2019.* Retrieved from https://www.educationcounts.govt.nz/__data/assets/pdf_file/0011/196562/Attendance-at-licensed-ECE-services-in-2019.pdf

Ministry of Research Science and Technology. (2007). Vision Mātauranga. Unlocking the innovation potential of Māori knowledge, resources and people. In: MoRST Wellington.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the Natrional Academy of Sciences USA, 108*(7), 2693-2698.

Moorfield, J. C. (2019). *Te Aka Online Māori Dictionary*.

Morton, S., Grant, C., Berry, S. D., Walker, C. G., Corkin, M., Ly, K., . . . Bandara, D. K. (2017). Growing Up in New Zealand: A longitudinal study of New Zealand children and their families. Now We Are Four: Describing the preschool years. Retrieved from http://ro.uow.edu.au/sspapers/3130

Morton, S., Grant, C. C., Walker, C. G., Berry, S. D., Meissel, K., Ly, K., . . . Kim, H. (2018). *Growing Up in New Zealand: A longitudinal study of New Zealand children and their families: Transition to school*. http://www.growingup.co.nz/sites/growingup.co.nz/files/2019-10/Report08_Transition%20to%20school%20June%202018.pdf

Munford, R., Sanders, J., Maden, B., & Maden, E. (2007). Blending whanau/family development, parent support and early childhood education programmes. *Social Policy Journal of New Zealand, 32*, 72-87. Retrieved from ww.minedu.govt.nz/index.cfm?layout=index&indexid=10987&indexparentid=1094

Murray, L., Jennings, S., Mortimer, A., Prout, A., Melhuish, E., Hughes, C., . . . Cooper, P. J. (2018). The impact of early-years provision in Children's Centres (EPICC) on child cognitive and socio-emotional development: study protocol for a randomised controlled trial. *Trials, 19*(1), 450. doi:10.1186/s13063-018-2700-x

Nakata, M. (2016). Indigenous knowledge and the cultural interface: Underlying issues at the intersection of knowledge and information systems. *IFLA Journal, 28*(5-6), 281-291. doi:10.1177/034003520202800513

Neha, T. (2016). *He maumahara ki nga korero a nga whanau ki Te Waipounamu: Family recollections and social contributions to Maori children's learning in the South Island, New Zealand.* (PhD). University of Otago, Retrieved from http://hdl.handle.net/10523/6341

Nix, R. L., Bierman, K. L., Domitrovich, C. E., & Gill, S. (2013). Promoting children's social-emotional skills in preschool can enhance academic and behavioral functioning in kindergarten: Findings from Head Start REDI. *Early Education & Development, 24*(7), 1000-1019.

Noddings, N. (2012). The caring relation in teaching. *Oxford Review of Education, 38*(6), 771-781. doi:10.1080/03054985.2012.745047

Peterson, E. R., Dando, E., D'Souza, S., Waldie, K. E., Carr, A. E., Mohal, J., & Morton, S. M. (2018). Can

Infant Temperament Be Used to Predict Which Toddlers Are Likely to Have Increased Emotional and Behavioral Problems? *Early Education and Development, 29*(4), 435-449. doi:10.1080/10409289.2018.1457391

Poulton, R., Moffitt, T. E., & Silva, P. A. (2015). The Dunedin Multidisciplinary Health and Development Study: overview of the first 40 years, with an eye to the future. *Social Psychiatry and Psychiatric Epidemiology, 50*(5), 679-693. doi:10.1007/s00127-015-1048-8

Rameka, L. (2011). Being māori: Culturally relevant assessment in early childhood education. *Early Years, 31*(3), 245-256. doi:10.1080/09575146.2011.614222

Rameka, L. (2012). Whakapapa: Culturally valid assessment in early childhood. *Early Childhood Folio, 16*(2), 33. Retrieved from https://search.informit.com.au/documentSummary;dn=112061245161691;res=IELHSS

Ratima, M., & Jenkins, B. (2012). *Whānau ora health needs assessment - Māori living in Taranaki*. New Plymouth: Taranaki District Health Board Retrieved from http://www.tdhb.org.nz/misc/documents/WOHNA_Full_%20Report_2012.pdf

Ratima, M., Theodore, R., Tamati, A., Hond-Flavell, E., Edwards, W., Korewha, H., . . . Poulton, R. (2019). Te Kura Mai i Tawhiti Research Programme: A collaborative lifecourse approach to health, wellbeing and whānau development. *MAI Journal, 8,* 63-76. doi:10.20507/MAIJournal.2019.8.1.5

Rauika Māngai. (2020). *A Guide to Vision Mātauranga: Lessons from Māori Voices in the New Zealand Science Sector*. Wellington, NZ: : Rauika Māngai Retrieved from https://ourlandandwater.nz/wp-content/uploads/2020/06/Rauika-M%C4%81ngai_A-Guide-to-Vision-M%C4%81tauranga_FINAL.pdf

Ritchie, J. (2008). Honouring māori subjectivities within early childhood education in aotearoa. *Contemporary Issues in Early Childhood, 9*(3), 202-210. doi:10.2304/ciec.2008.9.3.202

Santos, A. J., Vaughn, B. E., Peceguina, I., Daniel, J. R., & Shin, N. (2014). Growth of social competence during the preschool years: A 3-year longitudinal study. *Child Development, 85*(5), 2062-2073. doi:10.1111/cdev.12246

Schimmel, L. (2018). *A Meta-analysis of Academic and Social/Emotional Benefits Associated with Preschool Participation.* (10811142 Psy.D.). University of Colorado at Denver, Ann Arbor. ProQuest Dissertations & Theses A&I database.

Schindler, H. S., Kholoptseva, J., Oh, S. S., Yoshikawa, H., Duncan, G. J., Magnuson, K. A., & Shonkoff, J. P. (2015). Maximizing the potential of early childhood education to prevent externalizing behavior problems: A meta-analysis. *Journal of School Psychology, 53*(3), 243-263. doi:10.1016/j.jsp.2015.04.001

Sharp, C., Croudace, T. J., Goodyer, I. M., & Amtmann, D. (2005). The Strengths and Difficulties Questionnaire: Predictive validity of parent and teacher ratings for help-seeking behaviour over one year. *Educational and Child Psychology, 22*(3), 28. Retrieved from https://uh.edu/class/psychology/clinical-psych/_docs/SharpSDQ_2005_crop.pdf

Sibley, C. G., & Houkamau, C. A. (2013). The Multi-Dimensional Model of Maori Identity and Cultural Engagement: Item Response Theory Analysis of Scale Properties. *Cultural Diversity & Ethnic Minority Psychology, 19*(1), 97-110. doi:10.1037/a0031113

Smith, G. H., & Smith, L. T. (2018). Doing Indigenous work: decolonizing and transforming the academy. *Handbook of Indigenous Education*, 1-27. doi:10.1007/978-981-10-1839-8_69-1

Sturrock, F., Gray, D., Fergusson, D., Horwood, J., & Smits, C. (2014). Incredible Years Follow-up Study–Long-term follow-up of the New Zealand Incredible Years Pilot Study. *Ministry of Social Development*. Retrieved from http://www.incredibleyears.com/wp-content/uploads/indredible-years-follow-up-study-new-zealand-2014.pdf

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). The final report: Effective pre-school education. *London: Institute of Education*. Retrieved from https://discovery.ucl.ac.uk/id/eprint/10005308/1/EPPE12Sylva2004Effective.pdf

Tamati, A., Hond-Flavell, E., & Korewha, H. (2008). Te Kōpae Piripono Centre of Innovation Research Report. In *Education Counts*. Wellington: Ministry of Education.

Tamati, A., Ratima, M., Hond-Flavell, E., Edwards, W., Hond, R., Korewha, H., Theodore, R., Treharne, G.J. & Poulton, R. (2021a). He Piki Raukura : Understanding Strengths-based Māori Child Development Constructs in Kaupapa Māori Early Years Provision. *MAI Journal, 10*(1), 17-29. http://www.journal.mai.ac.nz/sites/default/files/MAI_Jrnl%202021_V10_I1_Tamati_FINAL.pdf

Tamati, A., Treharne, G. J., Theodore, R., Ratima, M., Hond-Flavell, E., Edwards, W., Hond, R. & Poulton, R. (2021b). He Piki Raukura: Assessing Ao Māori developmental constructs – Part I: Reliability of novel strengths-based measures among preschool Māori children. *New Zealand Journal of Psychology, 50*(2), 22-34.

Taylor, Z. E., Eisenberg, N., Spinrad, T. L., Eggum, N. D., & Sulik, M. J. (2013). The relations of ego-resiliency and emotion socialization to the development of empathy and prosocial behavior across early childhood. *Emotion, 13*(5), 822. doi:10.1037/a0032894

Theodore, R., Ratima, M., Edwards, W., Sporle, A., Te Morenga, L., Kiro, C., & Hond, R. (2019). How a lifecourse approach can promote long-term health and wellbeing outcomes for Māori. *Journal of Indigenous Wellbeing, 4*(1), 15-25. Retrieved from https://journalindigenouswellbeing.com/media/2019/07/122.126.How-a-lifecourse-approach-can-promote-long-term-health-and-wellbeing-outcomes-for-M%C4%81ori.pdf

Venker, C. E., Ray-Subramanian, C. E., Bolt, D. M., & Weismer, S. E. (2014). Trajectories of autism severity in early childhood. *Journal of autism and Developmental Disorders, 44*(3), 546-563.

Watts, T. W., Gandhi, J., Ibrahim, D. A., Masucci, M. D., & Raver, C. C. (2018). The Chicago School Readiness Project: examining the long-term impacts of an early childhood intervention. *PLoS One, 13*(7). doi:10.1371/journal.pone.0200144

Williams, H. W. (1957). *A dictionary of the Māori language*. Retrieved from http://nzetc.victoria.ac.nz/tm/scholarly/tei-WillDict.html

Williams, L. R., Degnan, K. A., Perez-Edgar, K. E., Henderson, H. A., Rubin, K. H., Pine, D. S., . . . Fox, N. A. (2009). Impact of behavioral inhibition and parenting style on internalizing and externalizing problems from early childhood through adolescence. *Journal of Abnormal Child Psychology, 37*(8), 1063-1075. doi:10.1007/s10802-009-9331-3

Wilson, W. H., & Kamana, K. (2001). Mai loko mai O ka 'I'ini: Proceeding from a dream. *The green book of language revitalization in practice, 147*.

Wylie, C., & Thompson, J. (2003). The Long-term contribution of early childhood education to children's performance--evidence from New Zealand. *International Journal of Early Years Education, 11*(1), 69-78. doi:10.1080/0966976032000066109

Yew, S. G. K., & O'Kearney, R. (2015). The role of early language difficulties in the trajectories of conduct problems across childhood. *Journal of Abnormal Child Psychology, 43*(8), 1515-1527. doi:10.1007/s10802-015-0040-9

Zhang, Z., McArdle, J. J., & Nesselroade, J. R. (2012). Growth rate models: Emphasizing growth rate analysis through growth curve modeling. *Journal of Applied Statistics, 39*(6), 1241-1262. doi:10.1080/02664763.2011.644528

Affiliations

Aroaro Tamati:
Taranaki Iwi, Ngāti Ruanui, Te Atiawa, Te Whānau-ā-Apanui

Jesse Kokaua:
Ngāti Ārera - Rarotonga

Mihi Ratima:
Ngāti Awa, Whakatōhea

Erana Hond-Flavell:
Taranaki Iwi, Ngāti Ruanui, Te Atiawa, Te Whānau-ā-Apanui

Will Edwards:
Taranaki Iwi, Ngāruahine, Tāngahoe, Pakakohi, Ngāti Ruanui

Ruakere Hond:
Taranaki Iwi, Ngāti Ruanui, Te Atiawa, Te Whānau-ā-Apanui

Reremoana Theodore:
Ngāpuhi

Corresponding Author

Aroaro Tamati
Te Pou Tiringa Incorporated
PO Box 6016, Moturoa
New Plymouth 4344
Aotearoa/New Zealand
Email: aroaro.tamati@gmail.com

Acknowledgements

We would like to sincerely thank the whānau of Te Kōpae Piripono and members of our Oversight Group who include Professor Sue Crengle, Professor Mera Penehira and Professor Leonie Pihama. We acknowledge funding for this research received from the Health Research Council of New Zealand in the form of a Ngā Kanohi Kitea Full Project Grant [Grant 13/954] and a Feasibility Grant [Grant 16/587]; in addition Reremoana Theodore was supported by a Māori Health Emerging Leader Fellowship [Grant 18/644]; Mihi Ratima was supported by a Ngā Pou Senior Fellowship [Grant 16/440]; and Ruakere Hond was supported by a Hōhua Tūtengaehe Research Fellowship [Grant 16/586]. We would also like to acknowledge the following who have provided funding for this research: the Ministry of Education, TOI Foundation (formerly TSB Community Trust), the University of Otago, the New Zealand Council for Educational Research, the National Centre for Lifecourse Research, and Te Pou Tiringa Inc.

Māori Glossary

Ao Māori	Māori world; Māori worldview
Aotearoa	Indigenous name for New Zealand
He Piki Raukura	one of the projects of Te Kōpae Piripono's longitudinal research, that focusing Māori child behavioural constructs
Kaitiaki	teacher at Te Kōpae Piripono
Kaupapa	purpose, objective, topic, philosophy
Kaupapa Māori	a Māori philosophical framework
Kōpae	shortened name of Te Kōpae Piripono (Taranaki-based Māori immersion early childhood centre)
Manawaroa	the notion of having courage in adversity, persisting despite difficulty and a positive outlook
Mātauranga Māori	Māori Indigenous knowledge systems
Piripono	the notion of having integrity, commitment and responsibility for a shared kaupapa/purpose
Tamariki	children
Tamariki Māori	Māori children
Taranaki	a tribe; a region in the west of the North Island
Te Kōpae Piripono	Taranaki-based Māori immersion early years and whānau initiative
Te Kura Mai i Tawhiti	the name given to Te Kōpae Piripono's longitudinal research programme
Te Pou Tiringa	Governing board of Te Kōpae Piripono
Tuakiri	the notion of a secure local Māori identity
Whai wāhi	inclusion
Whānau	family, usually encompassing wider membership than the nuclear family
Whānauranga	the notion of feeling and acting, as a member of a whānau/community

Residual Rape Myth Acceptance among Young Women Who Have Recently Completed a Sexual Violence Prevention Workshop

Madeline Hayward[1], Gareth J. Treharne[1], Nicola Liebergreen[2], Katie Graham[3] and Melanie Beres[2,3]
[1] Department of Psychology, University of Otago, Aotearoa/New Zealand
[2] Sociology, Gender Studies and Criminology, University of Otago, Aotearoa/New Zealand
[3] Te Whare Tāwharau, University of Otago, Aotearoa/New Zealand

This study explored young women's conceptions of sexual violence after attending a prevention workshop at university and addressed how rape myths feature in ongoing thinking about sexual violence. Three focus groups were carried out with a total of seven 18-/19-year-old women living in residential colleges who had recently attended a sexual violence prevention workshop during their first year at a university in Aotearoa/New Zealand. The focus groups involved activities to discuss rape myths and wider perspectives about tackling sexual violence on campus. Thematic analysis led to the development of three themes: women's lived experience of rape culture, women's residual rape myth acceptance, and encouraging men to challenge rape culture. These results demonstrate how rape myth acceptance can continue after attending a sexual violence prevention workshop and suggest that workshops should further address rape myths using evidence about how some such myths may be unintentionally reinforced.

Keywords: *Sexual violence prevention; Health education; Rape myths; Rape culture; Qualitative research; Focus groups*

Introduction

The prevalence of sexual violence makes the necessity of an intervention clear. In Aotearoa/New Zealand, approximately a third of women experience interpersonal or sexual violence at some point in their lives (Fanslow & Robinson, 2011). Rates of sexual violence on university campuses are especially high, which illustrates a need for sexual violence prevention programmes in this setting (Towl, 2018). However, the university sector currently lacks a comprehensive approach to sexual violence prevention (Beres, Stojanov, & Treharne, 2019; End Rape on Campus Australia, 2017). It has also been noted that universities have not until recently started challenging the societal norms that contribute to the prevalence and acceptance of sexual violence (End Rape on Campus Australia, 2017), and further research is needed to understand the status quo and inform change (Beres et al., 2019).

Sexual violence prevention workshops, such as bystander programmes, provide universities with an opportunity to demonstrate their commitment to reducing sexual assault on campuses (Banyard, Moynihan, & Plante, 2007; Beres et al., 2019). These prevention workshops are typically evaluated using quantitative methods. Qualitative research has the potential to expand on these quantitative evaluations by providing novel insights into how concepts such as rape myths function and thus inform understandings of sexual violence and rape myths in the university environment, where such workshops are increasingly common. This was achieved in the present study by considering how participants reflect on whether attending a sexual violence prevention workshops directly challenges rape myths or subtly reinscribes any such myths.

This study is an exploration of rape myths that was conducted alongside a larger project on the feasibility of a bystander sexual violence prevention workshop in Aotearoa/New Zealand. Bystander sexual violence prevention workshops seeks to counter sexual violence by teaching participants how to safely intervene in situations where they are bystanders (Banyard et al., 2007). Banyard, Eckstein, and Moynihan (2010) described how bystander workshops involve identifying situations where intervention is necessary and discussing strategies for safely intervening before, during, or after sexual violence. Participants also receive education about different types of sexual violence and their prevalence, the role of a bystander, and psychological findings about bystander behaviour. Throughout bystander programmes, participants are encouraged to consider and challenge societal contributions to the prevalence of sexual violence, such as rape myths and gender stereotypes around sex and sexual violence. Therefore, participants of such workshops are in a unique position to demonstrate how their understandings of rape myths become clarified or shift in light of discussions spurred during and after the workshop.

Bystander sexual violence prevention workshops are usually delivered to single-gender groups because rape prevention literature suggests that this method is more effective (Breitenbecher, 2000). Bystander workshops have been shown to have a number of positive outcomes for participants, including decreased rape myth acceptance and increased positive bystander behaviour after participation (Banyard et al., 2007; Moynihan, Banyard, Arnold, Eckstein, & Stapleton, 2010, 2011; Potter & Moynihan, 2011). Several studies have also shown that both empathy and education have a role to play in increasing bystanders' willingness to intervene (e.g., McMahon, 2010; Stewart, 2014).

The main limitation of the existing body of research on sexual violence prevention is the relative absence of studies applying qualitative methods to understand changes in key concepts such as rape myths from the perspective of participants. Previous evaluative studies of bystander sexual violence prevention programmes have involved quantitative methods demonstrating changes in the intended outcomes around bystanding (Banyard et al., 2007; Moynihan et al., 2010, 2011, 2015; Potter & Moynihan, 2011). Qualitative methods allow for exploration of how conceptualisations of rape myths change in response to structured programmes that address sexual violence. The present study utilised focus groups as a method for gaining rich exploratory data to explore how participants talk about rape myths within a social group with shared experience of a bystander sexual violence prevention workshop.

Rape myths are attitudes and beliefs about sexual violence that are contrary to established data. For example, a common misconception is that most instances of sexual violence occur between strangers (End Rape on Campus Australia, 2017; Gavey, 2019), when evidence suggests about 90% of sexual assaults in Aotearoa/New Zealand occur between people who are known to each other (Morris, Reilly, Berry, & Ransom, 2003). This statistic reflects an international trend where, in most instances of rape or sexual assault, the perpetrator is known to the victim (Ben-David & Schneider, 2005; End Rape on Campus Australia, 2017; Kelleher & McGilloway, 2009; McMahon, 2010).

Rape myths are one aspect of rape culture. Rape culture is a phenomenon where a community holds views which normalise and minimise the impact of sexual assault; in such a community, sexual assault becomes dismissible (Gavey, 2019). Rape myths contribute to rape culture by silencing and dismissing claims of rape, and by shifting the burden of responsibility for rape from perpetrators and onto victims (Giraldi & Monk-Turner, 2017). The latter phenomenon is known as victim blaming (Singleton, Winskell, Nkambule-Vilakatib, & Sabben, 2018). A classic example of victim blaming is the notion that it is a woman's fault if she gets sexually assaulted when she is dressed a certain way (Payne, Lonsway, & Fitzgerald, 1999). Other common rape myths include the misconceptions that women 'ask for' sexual assault through their conduct, that rapists are not accountable for their actions, that forced sex cannot be considered rape if the victim did not protest, and that victims of sexual assault are merely lying (Payne et al., 1999).

Qualitative data highlights the detrimental role of victim blaming on sexual assault survivors' wellbeing and access to support services. Kelleher and McGilloway (2009) explored the perceptions of service providers in the sexual violence sector around barriers to accessing care after sexual assault. Participants spoke about survivors feeling shame and guilt due to the belief that the assault was their fault. Further, participants spoke about survivors who had experienced negative reactions when they disclosed assault, due to their confidante reinforcing the notion of victim responsibility. Survivors' guilt and shame was a significant barrier to disclosing the incident again, often preventing the incident from being reported to the police, and a significant barrier in accessing support (Kelleher & McGilloway, 2009).

Petersen, Bhana and McKay (2005) used focus groups to explore how young women and men in South Africa spoke about the risks of becoming victims or perpetrators of sexual violence. Both female and male participants spoke about the notion of male superiority as a factor used to justify rape, and about rape being used by men to dominate women. Young men spoke about being pressured to perpetrate rape in order to prove their masculinity. Participants explained that gender-based violence is a norm in their community. Both female and male participants also spoke about rape myths as legitimising sexual assault, thus showing rape myth acceptance. The themes of Petersen et al.'s (2005) study help to demonstrate the role of rape myths in maintaining tolerance of sexual violence.

Lower levels of rape myth acceptance are associated with more instances of pro-social bystander behaviour and attitudes (Banyard, 2008), and higher levels of rape myth acceptance are associated with less willingness to intervene as a bystander (McMahon, 2010). Higher rape myth acceptance among men is associated with more hostile attitudes and behaviours towards women (Suarez & Gadalla, 2010) and is theorised to be a precursor to perpetrating sexual assault (Russell & King, 2016). Higher rape myth acceptance is associated with having no previous education about rape prevention (McMahon, 2010), which points to a role for education in lowering rape myth acceptance and thereby increasing pro-social bystander behaviour.

The aim of this study was to explore the link between rape myth rejection and participation in sexual violence prevention workshops. The present study used focus groups comprised of young women who had recently completed a bystander sexual violence prevention workshop. Considering that bystander workshops are well evaluated, the intention here was not to qualitatively evaluate the programme. It is already established that bystander interventions are effective at reducing rape myth acceptance (Banyard et al., 2007; Moynihan et al., 2010, 2011, 2015; Potter & Moynihan, 2011), and the purpose of this study was to explore in more detail how young women make sense of rape culture after such a workshop. There were two research questions: How do young women who have recently participated in a bystander sexual violence prevention workshop conceptualise rape culture and rape myths? Do young women indicate that participating in a sexual violence

prevention workshop affected their understanding of rape culture and rape myths?

METHODS

Design

This qualitative study involved three semi-structured focus groups. Focus groups were used to meet the aim of exploring how the social issue of rape myths was talked about in a social setting. Braun and Clarke's (2006) approach to thematic analysis was used to organise the qualitative data into themes. Themes were identified at a semantic level following a theory-driven process in relation to the concept of rape myths. A realist position informed the analysis. The data used in this analysis were gathered as part of a larger study of bystander sexual violence prevention workshops and this paper only focuses on participants' discussion of rape culture and rape myths, especially in relation to participation in the workshop. Focussing on one aspect of a qualitative dataset is an accepted approach to thematic analysis (Braun & Clarke, 2006). The overall study and this qualitative component received approval from the Human Ethics Committee of the University of Otago, where the research was conducted (Stojanov et al., 2021).

Participants

Seven participants were recruited from two residential colleges at the University of Otago in Dunedin, Aotearoa/New Zealand. All participants had completed a bystander sexual violence prevention workshop at their college in the previous month (Stojanov et al., 2021). One participant had completed the workshop sessions with a mixed group of young men and women; the other six had completed sessions with other young women only. Five participants were 18 years old and two participants were 19 years old. All participants identified as female. Four participants identified as Pākehā/New Zealand European, one participant as Aboriginal, one participant as Asian, and one participant as Scottish. Five participants identified their sexuality as straight and two participants identified as bisexual/pansexual. Four participants were studying psychology, with the rest studying a variety of other humanities and science subjects. Each focus group consisted of participants from the same residential college.

Materials

A semi-structured schedule was created for the focus groups by two researchers involved with this study and is available on request. The schedule was designed to be flexible and guided by participants' responses, so as to ensure that the resulting data were rich and reflected participants' views as best as possible. The schedule contained questions exploring participants' understandings of rape myths and rape culture as well as their experiences of the bystander sexual violence prevention workshop. The questions were mostly open-ended, with the intention of eliciting detailed answers (Braun & Clarke, 2006).

During the focus groups, several materials were used to help elicit responses. An activity using five pairs of cards was adapted from our previous research (Graham et al., 2021). One card in each pair had a question about a rape myth and the corresponding card had a statistic or statement that was counter to the myth. For example, one card read 'a victim will always scream, fight and act hysterical if someone tries to rape them. True or false?' Its pair read 'false' (Payne et al., 1999). The other four cards pertained to the rape myths that more instances of sexual assault occur between strangers, that alcohol causes sexual assault, that unwanted intercourse cannot be called rape if physical force was not used (End Rape on Campus Australia, 2017; Payne et al., 1999), and that perpetrators of sexual violence are mentally ill (Cowan & Quinton, 2006). Most of these myths were related to those within the Illinois Rape Myth Acceptance Scale (Payne et al., 1999), which all participants had completed prior to participation in the bystander sexual violence prevention workshop as part of the quantitative evaluation in the larger study.

Two images that had been presented and discussed during the bystander sexual violence prevention workshops were adapted for use in the focus groups. Each image showed a group of people, with speech bubbles indicating whether they were conveying acceptance of, or disagreeing with, a rape myth. One of these images related to a myth about the role of alcohol in sexual assault, and the other related to the myth that a woman's clothing is related to sexual assault (Payne et al., 1999).

Photographs taken around the university campus and its surrounding area were printed out and used to facilitate discussion about the type of scenario where participants thought sexual assault would be most likely to occur. This prompt related to the rape myth that sexual assault is something that mainly occurs in seedy locations (Payne et al., 1999). The statements "Sex workers can't be raped" and "Men can't be raped because they always want sex" were printed on cards and presented to participants to generate discussion about these rape myths.

Procedure

All participants who had completed the bystander sexual violence prevention workshops were invited by email to attend a focus group. The focus groups were conducted by two female researchers (the first author and third author). A third female researcher observed each session and took notes. The duration of focus groups ranged from just over one hour to just over an hour and a half. The first focus group, comprising three participants, was held in a private study room at the participants' college. The other two, each comprising two participants, were held in a quiet, private meeting room on campus. Each focus group was audio-recorded and later transcribed by a professional service. All participants gave informed consent before the focus group started and were given a $15 supermarket voucher as reimbursement for expenses related to participating.

The focus groups began with introductions and questions about participating in the bystander sexual violence prevention workshops. Next, a facilitator used the materials described above to generate discussion about rape myths, beginning with the task involving five questions about rape myths. Following a discussion about this task, the other prompts were presented. Generally, after a prompt was presented, participants were asked a broad, open question, such as, "What do you think these pictures represent?" There was a particular focus on generating further discussion around any myths that

appeared to be contentious or difficult for participants to respond to.

Analysis

The transcripts were checked for accuracy and to ensure they were anonymised. This process, along with repeated readings of the transcripts, lead to familiarisation with the data and formed the first stage of thematic analysis (Braun & Clarke, 2006). Analysis was led by the first author and discussed with the other authors. In line with Braun and Clarke's (2006) guidelines for thematic analysis, the next stage of analysis involved systematically generating codes from all focus group data. These codes were then organised into potential themes that represent repeated patterns in the data pertaining to the research questions. Themes were then reviewed in relation to their associated coded extracts as well as the entire data set. Once themes had been checked, they were named and defined following the steps described by Braun and Clarke (2006).

Themes were identified at a semantic level, with the main interest being what was described by participants, rather than theorising underlying ideas or assumptions (Braun & Clarke, 2006). A semantic approach to analysis was in line with our application of a realist ontology/epistemology and enabled us to take a broad approach to exploring themes that represent an analysis of the perspectives expressed by the young women who participated (Braun & Clarke, 2006). Given that the focus group schedule and the analysis were driven by a theoretical interest in rape myths, the process of identifying themes was theory-driven (Braun & Clarke, 2006). Quotes were used to illustrate themes. In presenting quotes, the notation '[…]' has been used to indicate where part of a quote has been redacted for brevity without changing the intended meaning.

The analysis was informed by the epistemological position of realism (Braun & Clarke, 2006), theorising that the way participants talked about rape myths meaningfully revealed the extent to which they accepted or rejected the rape myths. In other words, the relationship between their discussion and the associated meaning was assumed to be a straightforward representation of experienced reality.

RESULTS

The thematic analysis of focus group data led to the development of three themes: 1) women's lived experience of rape culture, 2) women's residual rape myth acceptance, and 3) encouraging men to challenge rape culture.

Theme 1: Women's Lived Experience of Rape Culture

The first theme is about the ways in which women conveyed their lived experience of rape culture. Participants spoke about experiencing the harmful societal attitudes that normalise sexual assault and form rape culture (Gavey, 2019). Participants unequivocally believed rape culture existed in Aotearoa/New Zealand and expressed why it is important to challenge rape culture and the experience they had doing this.

It was common among participants to know people who had been sexually assaulted and many had themselves been the target of sexist behaviour. This experience was a motivating factor for participating in the workshop and shaped their perspectives on rape myths. With regard to the rape myth about alcohol causing sexual assault, one participant responded, *'I don't want to blame alcohol […] because a lot of the sexual assault I […] know about […] has nothing to do with alcohol'*. Participants' experience with rape culture also facilitated their understanding that it is *'the little things that sort of add up that make people think it's ok to rape people so like um rape jokes, sexist comments…'*.

Participants argued that someone who believed rape culture did not exist within Aotearoa/New Zealand would be *'incredibly naïve'*. They were also discerning as to how prominent aspects of rape culture known to exist within the US might be found within Aotearoa/New Zealand because, for example, *'we don't have fraternities here but we do have sports teams'*.

All participants thought it was important to challenge rape myths and many had experience of challenging rape myth acceptance and sexist attitudes in those around them. This was despite the liability that *'you'll get called names or like oh you can't take a joke'*. Hearing sexist views from family and friends was distressing to them: *'it like hurt me to think that someone that I'm quite close to as a friend would think these things'*.

Participants also noted that one of the most valuable parts of the workshop was learning about the *'spectrum'* of sexual violence ranging from jokes up to more harmful aspects. One participant reflected that *'some people don't understand just how bad the issue is and how little things like […] sex jokes or like misogynistic jokes […] how often they actually occur and […] in regards to the scale of sexual harassment, even though it's so little it happens so much more often than the big things'*. Rape jokes were therefore seen as important to challenge, especially after participating in the workshop: *'it sort of gave me a way to learn how to deal with these things because they happen so often and nobody sort of sticks up for it and it just becomes acceptable'*.

Theme 2: Women's Residual Rape Myth Acceptance

The second theme is about a subtle residual level of rape myth acceptance that related to a varied effect of the workshop on these myths. Most participants indicated that they already strongly rejected rape myths, with the workshop only strengthening their views. Other participants pointed to specific myths where participation in the workshop had changed their opinion by strengthening their rape myth rejection. However, some rape myth acceptance, both explicit and subtle, remained after participation in the workshop.

In most instances, the workshop mainly served to strengthen existing anti-rape myth views. In the words of one participant: *'I've always sort of had similar perspectives but I think that the course has definitely sort of cemented it and given me more reasons to believe it'*. This sentiment was also expressed in an exchange between two participants:

> Participant 1: *'I think personally I had pretty solid opinions anyway just like from the way I've been brought up and all that and the way I myself have*

learned about things […] you might've guessed but I have very solid opinions (laughs)'
Participant 2: *'Yeah same'.*
Participant 1: *'[…] so I think personally […] a lot of my beliefs were reaffirmed I guess […] but I wouldn't say they were changed'.*

However, in some cases, participants credited the workshop as forming or changing their understanding. As a case in point, several participants had thought that most instances of sexual assault were committed by strangers before the workshop. After the workshop, almost all participants acknowledged that evidence shows more instances of sexual assault occur between acquaintances. One participant said, *'I almost trust guys less […] because now I know that it's like it happens between friends and acquaintances'*. While instilling a distrust of men is not an intended outcome of the workshop, this response does show rejection of the myth that most rapes are committed by strangers.

Participants were generally quick to reject rape myths that were raised, providing lengthy, nuanced reasoning. However, even after participating in the workshop, participants did not unanimously disagree with the rape myth that 'perpetrators of sexual violence are mentally ill'. Instead, some participants argued that the statement could be true in the case of some mental illnesses because, for example, perpetrators might have a *'kind of complex in their head'* or be *'mentally ill as in someone who doesn't feel empathy'*. However, participants did not think a common mental illness such as depression or anxiety would be *'the sort of thing that would push you to do that [i.e., commit sexual violence]'*.

In other cases, participants explicitly rejected a rape myth, but their comments revealed subtle rape myth acceptance. Participants stated that the workshop had made them aware that sexual assault can *'literally happen anywhere'* and not just in *'dark alleys and stuff'*, thus challenging this rape myth. However, all participants indicated that they thought sexual assault would be most likely on a dark street at night, showing subtle residual acceptance of this myth: *'they're isolated, they're dark, there's not people around to hear you so it would be incredibly easy for someone to hurt you'*. Participants also implicitly indicated rape myth acceptance when reasoning that sexual assault would more likely occur in a large hall of residence where there was *'more anonymity'* despite explicitly rejecting the myth that more sexual assault is perpetrated by strangers: *'if there's more than 500 people [living in a hall], it's what you said about not knowing everyone and easier just to be another face in the crowd'*.

Theme 3: Encouraging Men to Challenge Rape Culture

The third theme is about how the female participants perceived men to be lacking a vested interest in challenging rape culture, and the ways this could be addressed. Participants perceived there to be a range of ways that men uphold rape culture. One was the way that men were seen to minimise women's reactions to sexist jokes, *'saying "Oh it's just a joke, why do you have to be so angry about everything?"'*. Participants also perceived a tendency among men to accept rape myths: *'they've got this idea well if she's drinking, then it's consent'*. Moreover, participants perceived that men's complicity in rape culture often means they are unintentionally abusive: *'they would just do whatever they want to get, like get, have sex […] they don't realise that there's like a whole rape part of it'*.

In addition, participants perceived men to be ignorant about sexual violence due to experiencing less *'general sexism'*, for example not being '*exposed to all of the stuff about clothes that we get*'. On the other hand, women were perceived to be more aware, due to a lived experience of rape culture: *'It's like with general sexism, we're more likely to notice it because it's addressed at us whereas guys in general are less likely to notice how sexist things are'*. One participant expanded on how they perceived women to be taught to be aware of sexual violence, saying *'they teach girls not to be raped and then there's less focus on guys not to rape so I feel like […] girls are more likely to get involved because they go well how can I stop this […] happening to me whereas guys kind of go well I'm not gonna join in with this because why would I, like I'm not gonna get affected'*. Participants noted that men are rarely active allies against sexual violence. They thought it was important that men should get involved: *'we were trying to make sure that guys went'* to the workshops.

Participants discussed three main reasons they perceived as preventing men from engaging with sexual violence prevention programmes. The first reason was the disapproval of other men: *'I think it's again that sort of idea of "Oh if I sign up, then what are other guys gonna think of me?"'*. The second reason was the idea that men do not think they will be affected by rape: *'they teach girls not to be raped and then there's less focus on guys not to rape […], girls are more likely to get involved because they go well how can I stop this happening […] whereas guys kind of go well I'm not gonna join in with […] I'm not gonna get affected*'. The third reason was the need to uphold hegemonic masculinity; that there is *'an idea of what a manly man is and […] if they invest time in this, then it sort of goes against this sort of idea'.* Participants suggested that enlightening men about issues of sexual violence could be done in a way *'that doesn't like offend their masculinity'*. Nonetheless, they thought it was important that hegemonic masculinity be challenged.

Participants acknowledged the common misconception that only women are affected by sexual assault, but explicitly rejected the myth that men cannot be raped: '*there are male victims as well'*. Participants acknowledged that anyone can perpetrate sexual assault or sexist behaviours. The need to uphold *'social stereotypes and gender roles'* was perceived to be detrimental to male victims of sexual assault: *'not a lot of guys actually come forward because there is that whole like sense of macho kind of stuff, like when you're a man, you don't get raped […] it's kind of the culture in guys for some reason and it's embarrassing to say that they've been raped […] so they don't come forward'*.

DISCUSSION

The aim of this study was to explore how young women make sense of rape culture and rape myths after participating in a bystander sexual violence prevention programme. Participants discussed rape culture as something they, or people close to them, had experienced.

Participants tended to reject rape myths, sometimes attributing this rejection to perspectives they had learned in the sexual violence prevention workshops. However, participants showed subtle rape myth acceptance after participation in the workshop. Participants spoke about a perceived lack of male engagement in sexual violence prevention and the ways this could be addressed.

Female participants spoke about their lived experience of rape culture. This finding should be of little surprise given the prevalence of sexual violence towards women. The way women discussed rape culture in the wake of the workshop implies that they were able to relate the workshop content to what they or others had experienced. Participants found the workshop valuable for helping them continue to challenge rape culture. Participants discussed the bystander sexual violence prevention workshop as having strengthened their rejection of rape myths. This finding aligns with previous literature that shows an association between participation in the bystander workshop and decreased rape myth acceptance (Moynihan et al., 2010). However, a reduction in rape myth acceptance does not mean a complete rejection of rape myths. Participants in this study explicitly accepted the myth that perpetrators of sexual assault are mentally ill and appeared to implicitly accept the myths that more sexual assaults occur between strangers or on dark streets at night.

Overall, these findings help identify subtle aspects of rape myths that may be more resistant to change than quantitative research has suggested. In this study, participation in a bystander sexual violence prevention workshop did seem to decrease participants' explicit rape myth acceptance for most myths, as has been found in past research (Banyard et al., 2007; Moynihan et al., 2010, 2011, 2015; Potter & Moynihan, 2011), but it may not have decreased implicit rape myth acceptance. Future research could investigate residual and implicit rape myth acceptance further and explore whether certain rape myths are more resistant to change than others. Qualitative methods made the exploration of nuanced perspectives on these myths possible. Open questioning in the focus groups revealed participants' implicit rape myth acceptance, whereas traditional survey measures may have only captured participants' explicit rejection of the myths. This finding highlights the necessity of ongoing qualitative and quantitative evaluations of sexual violence prevention programmes.

The female participants in this study perceived men to be lacking a vested interest in challenging rape culture. Participants perceived men to be oblivious to issues of sexual violence. Previous literature has suggested that women tend to have greater knowledge about sexual violence than men (Banyard, 2008). Participants perceived that men tend to have higher rape myth acceptance, and this is consistent with previous literature in which women have lower rape myth acceptance than men (Suarez & Gadalla, 2010).

Participants also had a perception that social expectations of hegemonic masculinity dissuaded men from getting involved with sexual violence prevention. This perception is consistent with literature exploring the link between masculinity and sexual violence. For example, the pressure to conform to masculine norms and the acceptance of behaviour aimed at objectifying women were found to potentially mediate the relationship between fraternity membership and acceptance of sexual violence in a sample of male American college students (Seabrook, Ward, & Giaccardi, 2018). Men's acceptance of and adherence to masculine stereotypes is thought to both justify the degradation of women and absolve men of responsibility for such behaviours (Giraldi & Monk-Turner, 2017). These findings imply that it is important to find ways to encourage men's involvement with sexual violence prevention. Participants discussed some ways that this could be achieved: educating men about sexual violence in a way that does not offend their masculinity and shifting norms among men so that engaging with sexual violence prevention is accepted.

The present study had several limitations, one being that the sample only included a moderate portion of the overall cohort. Given that seven women out of a larger group who completed the bystander sexual violence prevention workshop participated in focus groups, the themes arising from this sample should not necessarily be taken as being reflective of all participants in the wider study. Future research could use similar methods with a larger cohort and could also use similar methods with male participants.

Another limitation was that participants had volunteered for both the bystander sexual violence prevention workshop and the focus groups. Participants whose participation in a bystander workshop is compulsory might discuss rape myth acceptance in a different way. Future research could explore rape myth acceptance in participants of compulsory sexual violence prevention workshops. In our study, the total number of participants who elected to participate in the workshop sessions was only a small portion of those invited to attend. If universities were to implement sexual violence prevention workshops on a wider scale, it would be important to consider whether compulsory sessions would be successful. For example, it would be worth exploring whether compulsory workshop attendance is associated with unintended 'backlash' effects (see Moynihan et al., 2010, 2011), where participants' attitudes or behaviours worsen after the intervention.

Participants in this study were all in their first year at university. Banyard and Moynihan (2011) noted that first-year students are at a unique developmental stage, meaning that the way they respond to sexual violence prevention messaging may not be reflective of all students' responses. It would be beneficial to explore rape myth acceptance in students of different year groups who have participated in sexual violence prevention workshops.

An established aspect of bystander sexual violence prevention workshops is open discussion (Banyard et al., 2010). Participants who attended different sessions of the workshops may have been involved in quite different discussions despite the overall standardisation. The extent to which these discussions shaped participants' views is unknown. A related limitation is that the focus groups took place up to several weeks after participants had attended the bystander workshops. Other events in participants' lives during this time may have affected their perspectives on the issues raised in the bystander

workshops. This limitation is common in studies evaluating the bystander workshop (e.g., Moynihan et al., 2010). However, it is worth noting that participants in the present study did appear to remember content from the bystander workshops well, which is another positive outcome.

Sexual violence prevention efforts are needed to ameliorate the high rates of sexual violence in Aotearoa/New Zealand (Fanslow & Robinson, 2011), particularly the disproportionately high rates experienced by Māori (Ministry of Justice, 2015, 2019). It is therefore critical that prevention efforts like bystander workshops are culturally meaningful to Māori. Definitions of sexual violence are not universal and Pākehā definitions of sexual violence are not necessarily relevant to Māori (Pihama et al., 2016). Intergenerational trauma, forced migration, and the mandated adoption of Western ideologies through colonisation have been identified as major contributors to the disparity in rates of sexual violence between Māori and non-Māori New Zealanders (Cavino, 2016; Pihama et al., 2016; Robertson & Oulton, 2008). It is important to consider how sexual violence prevention efforts can include definitions of sexual violence that are located within a Māori worldview, for example by considering sexual violence as an act which harms both individual and collective well-being (Pihama et al., 2016).

The bystander sexual violence prevention workshop that our participants attended was originally developed in the US by Banyard and colleagues (Banyard et al., 2007; Moynihan et al., 2010, 2011, 2015; Potter & Moynihan, 2011). Translating this workshop to the cultural context of the participants involved incorporating discussions of media about local cases of sexual violence. Future consideration must be given to whether this modification alone is sufficient when delivering such workshops in the context of in Aotearoa/New Zealand. Future research involving Māori participants and led by Māori researchers working within a Kaupapa Māori framework may help inform an understanding of how these workshops could incorporate definitions of sexual violence beyond dominant Western understandings.

Furthermore, future research could explore whether the subtle aspects of rape myth acceptance and overt aspects of rape myth rejection discovered in this study remain for longer periods after attending sexual violence prevention workshops and how those aspects translate into everyday interactions, such as responding when rape jokes are made by others. Understanding the long-term trajectory of rape myth rejection would allow universities to make decisions about implementing workshops on a large scale and understanding whether it is sufficient for students to participate once during their time at university or whether booster sessions might be beneficial to address subtle rape myth acceptance.

In this study, semantic analysis led to the development of themes that broadly captured the perspectives expressed by the young women who participated. In future studies, particularly of larger data corpuses, it could be useful to take a latent approach to analysis (as outlined by Braun & Clarke, 2006, and others who have provided guidance for discursive approaches to analysis). Latent analyses would allow deeper exploration of how participants' experiences and reflections of sexual violence prevention workshops can be understood in terms of wider discourses about sexual violence, as this will be pertinent to informing how the workshops can meet their goal around enduring attitudinal change. For example, a latent approach could provide a deeper insight into findings of the present study such as the residual endorsement of the rape myth that sexual assault is more common on a dark street, by considering how this finding can be understood in relation to endorsement of victim blaming. Similarly, findings around women's perceptions of men's disinterest in sexual violence prevention could be more deeply understood in relation to wider discourses around masculinity and gender roles.

The current study only involved participants who had completed the bystander sexual violence prevention workshop. There was no exploration of whether different themes about rape myths arose in focus groups with participants who had not participated in a bystander sexual violence prevention workshop. Further, the extent to which participating in the bystander workshop affected participants' perspectives on rape myths could only be investigated by asking participants whether they believed they had held different attitudes before the workshop; there was no quantitated comparison of attitudes before and after participation in the workshop. These limitations suggest two directions for future research. Firstly, themes arising in focus groups using participants who either did or did not complete the programme could be compared. Secondly, focus groups could be carried out with one group of participants before and after they participated in a bystander sexual violence prevention workshop, thus allowing themes from each time point to be compared.

This qualitative study has provided novel insights into how young women who have completed a sexual violence prevention workshop discuss rape culture. The female participants spoke about a lived experience of rape culture and their perception that men need to be encouraged to challenge rape culture. The analysis identified some aspects of rape myth acceptance that remained after the intervention. The results imply that there is scope to investigate how certain aspects of sexual violence prevention workshops might be modified in order to optimise their intended outcomes.

References

Banyard, V. L. (2008). Measurement and correlates of prosocial bystander behavior: The case of interpersonal violence. *Violence and Victims, 23*(1), 83-97. https://doi.org/10.1891/0886-6708.23.1.83

Banyard, V. L., Eckstein, R. P., & Moynihan, M. M. (2010). Sexual violence prevention: The role of stages of change. *Journal of Interpersonal Violence, 25*(1), 111-135. https://doi.org/10.1177/0886260508329123

Banyard, V. L., & Moynihan, M. M. (2011). Variation in bystander behavior related to sexual and intimate partner violence prevention: Correlates in a sample of college students. *Psychology of Violence, 1*(4), 287-301. https://doi.org/10.1037/a0023544

Banyard, V. L., Moynihan, M. M., & Plante, E. G. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology, 35*(4), 463-481. https://doi.org/10.1002/jcop.20159

Ben-David, S., & Schneider, O. (2005). Rape perceptions, gender role attitudes, and victim-perpetrator acquaintance. *Sex Roles, 53*(5/6), 385-399. https://doi.org/10.1007/s11199-005-6761-4

Beres, M. A., Treharne, G. J., & Stojanov, Z. (2019). A whole campus approach to sexual violence: The University of Otago Model. *Journal of Higher Education Policy and Management, 41*(6), 646-662. https://doi.org/10.1080/1360080X.2019.1613298

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Breitenbecher, K. H. (2000). Sexual assault on college campuses: Is an ounce of prevention enough? *Applied and Preventive Psychology, 9*(1), 23–52. https://doi.org/10.1016/S0962-1849(05)80036-8

Cavino, H. M. (2016). Intergenerational sexual violence and whānau in Aotearoa/New Zealand - pedagogies of contextualisation and transformation. *Sexual Abuse in Australia and New Zealand: An Interdisciplinary Journal, 7*(1), 4-17. https://www.anzatsa.org/saanz-journal/browse-saanz-articles//show/65

Cowan, G., & Quinton, W. J. (2006). Cognitive style and attitudinal correlates of the perceived causes of rape scale. *Psychology of Women Quarterly, 21*(2), 227-245. https://doi.org/10.1111/j.1471-6402.1997.tb00110.x

End Rape on Campus Australia. (2017). *Connecting the dots: Understanding sexual assault in university communities* (submission to the Australian Human Rights Commission's 'University Sexual Assault and Harassment' project). Australia: End Rape on Campus Australia

Fanslow, J. L., & Robinson, E. M. (2011). Sticks, stones, or words? Counting the prevalence of different types of intimate partner violence reported by New Zealand women. *Journal of Aggression, Maltreatment & Trauma, 20*(7), 741-759. https://doi.org/10.1080/10926771.2011.608221

Gavey, N. (2019). *Just sex? The cultural scaffolding of rape* (2nd ed.). Abingdon: Routledge.

Giraldi, A., & Monk-Turner, E. (2017). Perception of rape culture on a college campus: A look at social media posts. *Women's Studies International Forum, 62,* 116-124. http://dx.doi.org/10.1016/j.wsif.2017.05.001

Graham, K., Treharne, G. J., Liebergreen, N., Stojanov, Z., Shaw, R., & Beres, M. A. (2021). A qualitative exploration of barriers to university students' willingness to attend sexual violence prevention workshops. *Sex Education, 21*(2), 148-160. https://doi.org/10.1080/14681811.2020.1772225

Kelleher, C., & McGilloway, S. (2009). 'Nobody ever chooses this ...': A qualitative study of service providers working in the sexual violence sector – key issues and challenges. *Health and Social Care in the Community, 17*(3), 295–303. https://doi.org/10.1111/j.1365-2524.2008.00834.x

Laner, M. R., Benin, M. H., & Ventrone, N. A. (2001). Bystander attitudes toward victims of violence: Who's worth helping? *Deviant Behaviour, 22*(1), 23-42. https://doi.org/10.1080/016396201750065793

McMahon, S. (2010). Rape myth beliefs and bystander attitudes among incoming college students. *Journal of American College Health, 59*(1), 3-11. http://dx.doi.org/10.1080/07448481.2010.483715

Ministry of Justice. (2015). *2014 New Zealand crime and safety survey: Main findings*. Retrieved from https://www.justice.govt.nz/assets/Documents/Publications/NZCASS-201602-Main-Findings-Report-Updated.pdf

Ministry of Justice. (2019). *New Zealand Crime and Victims Survey: List of Key Findings Cycle 1 (March – September 2018).* Retrieved from https://www.justice.govt.nz/assets/Documents/Publications/NZCVS-List-of-key-findings-Y1-fin.pdf

Morris, A., Reilly, J., Berry, S., & Ransom, R. (2003). *New Zealand National Survey of Crime Victims 2001* (report on the National Survey of Crime Victims). Wellington, New Zealand: Ministry of Justice.

Moynihan, M. M., Banyard, V. L., Arnold, J. S., Eckstein, R. P., & Stapleton, J. G. (2010). Engaging intercollegiate athletes in preventing and intervening in sexual and intimate partner violence. *Journal of American College Health, 59*(3), 197-204. https://doi.org/10.1080/07448481.2010.502195

Moynihan, M. M., Banyard, V. L., Arnold, J. S., Eckstein, R. P., & Stapleton, J. G. (2011). Sisterhood may be powerful for reducing sexual and intimate partner violence: An evaluation of the Bringing in the Bystander in-person program with sorority members. *Violence Against Women, 17*(6), 703-719. https://doi.org/10.1177/1077801211409726

Moynihan, M. M., Banyard, V. L., Cares, A. C., Potter, S. J., Williams, L. M., & Stapleton, J. G. (2015). Encouraging responses in sexual and relationship violence prevention: What program effects remain 1 year later? *Journal of Interpersonal Violence, 30*(1), 110-132. https://doi.org/10.1177/0886260514532719

Payne, D. L., Lonsway, K. A., & Fitzgerald, L. F. (1999). Rape myth acceptance: Exploration of its structure and its measurement using the Illinois Rape Myth Acceptance scale. *Journal of Research in Personality, 33*(1), 27-68. https://doi.org/10.1006/jrpe.1998.2238

Petersen, I., Bhana, A., & McKay, M. (2005). Sexual violence and youth in South Africa: The need for community-based prevention interventions. *Child Abuse & Neglect, 29*(11), 1233-1248. https://doi.org/10.1016/j.chiabu.2005.02.012

Pihama, L., Te Nana, R., Cameron, N., Smith, C., Reid, J., & Southey, K. (2016). Māori cultural definitions of sexual violence. *Sexual Abuse in Australia and New Zealand: An Interdisciplinary Journal, 7*(1), 43-51. https://www.anzatsa.org/saanz-journal/browse-saanz-articles//show/68

Potter, S. J., & Moynihan, M. M. (2011). Bringing in the Bystander in-person prevention program to a U.S. military installation: Results from a pilot study. *Military Medicine, 176*(8), 870-975. https://doi.org/10.7205/MILMED-D-10-00483

Robertson, N., & Oulton, H. (2008). *Sexual Violence: Raising the conversations. A literature review.* Hamilton, New Zealand: The University of Waikato. https://researchcommons.waikato.ac.nz/handle/10289/995

Russell, T. D., & King, A. R. (2016). Anxious, hostile, and sadistic: Maternal attachment and everyday sadism predict hostile masculine beliefs and male sexual violence. *Personality and Individual Differences, 99*, 340-345. http://dx.doi.org/10.1016/j.paid.2016.05.029

Seabrook, R. C., Ward, M., & Giaccardi, S. (2018). Why is fraternity membership associated with sexual assault? Exploring the roles of conformity to masculine norms, pressure to uphold masculinity, and objectification of women. *Psychology of Men & Masculinity, 19*(1), 3-13. http://dx.doi.org/10.1037/men0000076

Singleton, R., Winskell, K., Nkambule-Vilakati, S., & Sabben, G. Young Africans' social representations of rape in their HIV-related creative narratives, 2005–2014: Rape myths and alternative narratives. *Social Science & Medicine, 198,* 112-120. https://doi.org/10.1016/j.socscimed.2017.12.032

Stewart, A. L. (2014). The Men's Project: A sexual assault prevention program targeting college men. *Psychology of Men & Masculinity, 15*(4), 481-485. http://dx.doi.org/10.1037/a0033947

Stojanov, Z., Treharne, G. J., Graham, K., Liebergreen, N., Shaw, R., Hayward, M., & Beres, M. (2021). Pro-social bystander sexual violence prevention workshops for first year university students: Perspectives of students and staff of residential colleges. *Kotuitui: New Zealand Journal of Social Sciences Online*. https://doi.org/10.1080/1177083X.2021.1906712

Suarez, E., & Gadalla, T. M. (2010). Stop blaming the victim: A meta-analysis on rape myths. *Journal of Interpersonal Violence, 25*(11), 2010-2035. https://doi.org/10.1177/0886260509354503

Towl, G. (2018, January). Tackling sexual violence at universities. *The Psychologist, 31,* 36-30.

Corresponding Author

Corresponding Author
Gareth J. Treharne
Department of Psychology / Te Tari Whakamātau Hinekaro
University of Otago / Te Whare Wānanga o Otāgo
William James Building, Leith Walk
PO Box 56
Dunedin 9054
Aotearoa/New Zealand
Email: gareth.treharne@otago.ac.nz

Acknowledgments

This research was funded by a University of Otago Research Grant. We thank the participants for their contribution and staff of the residential colleges for facilitating recruitment. We also thank Rachel Shaw for assisting with the focus groups.

Measuring Racial Prejudice in New Zealand Pre-schoolers and testing an intervention to reduce the same using brief cross-race friendship picture books

Anita A. Azeem, John A. Hunter and Ted Ruffman
Department of Psychology, University of Otago, Aotearoa/New Zealand

This study with 88 children from Dunedin (M_{Age} = 5.18, *SD* = 1.32) aimed to measure racial prejudice, particularly that against Asians and Arab Muslims. Each participant was requested to complete two tasks to measure their explicit bias and one task to measure their implicit bias. Together, the results indicated that young children display a greater preference for friendship with own race and rate children from own race more positively than children from other races. Further, when these participants were tested again post-intervention (a month of reading picture books about cross-race friendships) they did not show any change thereby indicating that this prejudice not only develops early but is also fairly rigid. Additionally, children's implicit prejudice displayed a positive relation with parent's racism score indicating that the children may have learnt such attitudes from their parents.

Keywords: *Prejudice in NZ, Muslim Prejudice, Asian Prejudice, Racial Prejudice, Preschool*

Introduction

In January 2020, a woman in Dunedin reported that her 9-year-old daughter was verbally abused by some boys of the same age who called her a 'stinky Indian' and rubbed dog excrement all over her face at a local playground (Ayling, 2020). While an incident of this nature is shocking, it indicates a deeper problem, i.e. the prevalence of racial prejudice in children. New Zealand has generally maintained its position as a peace-loving nation governed by an empathic leader who spoke regarding the Muslim community that, "they are us" immediately after the 2019 Christchurch mosque attack. However, certain discriminatory acts indicate that perhaps the wider population may not resonate with the Prime Minister. In fact, recent scientific findings within the country indicate that prejudice against Muslims is substantially higher than towards any other religious group (Greaves et al., 2020; Wilson, 2019). Further, Asians are the least warmly rated ethnic group (Sibley & Ward, 2013) and one fourth of New Zealanders feel that Asians are the most discriminated group in the country (Human Rights Commission, 2010). Taken together, these findings indicate that racial prejudice and discrimination against Asians and Muslims in New Zealand exists and cannot be denied.

While most of the current studies in the country have focused on measuring attitudes amongst adults, little is known about whether the same trends follow in children. The purpose of the current study, therefore, was to fill that gap in literature and to understand how young children in NZ feel about ethnic and racial minorities. Our research had two main goals: First, to measure pre-schoolers prejudice against Asians and Arab Muslims and second, to see if such prejudice could be reduced by introducing cross-race friends via picture books.

Development of Racial Preference

Elaborate research in prejudice development has indicated that racial preference starts at an early age. For instance, Kelly et al. (2005) found that at 3 months, children start displaying a preference for their own race, although such a preference is not present at birth. Further, Bar-Haim et al. (2006) found that infants who were exposed to multiple races did not show an own-race preference. This finding is crucial in understanding that early exposure to multiple races may reduce own race preferences. By 9 months, infants are able to categorise own-race faces into one category and other-race faces into a separate category (Quinn, et al., 2016), and begin to associate 'own race' with happy music and 'other race' with sad music (Xiao et al., 2017). Between 10 and 12 months, children show a preference towards snacks and toys endorsed by someone who speaks their native language (Shutts et al., 2009) and between 14 and 18 months, toddlers imitate physical actions of a native speaker more than those of a foreigner (Buttelmann et al., 2013).

The same trends continue in pre-schoolers (age 2 to 5) who prefer to play with and allocate more resources to people from their 'own race' (Kinzler & Spelke, 2011; Renno & Shutts, 2015) particularly when distributing limited resources (Lee et al., 2018). A similar pattern was also reported by Fehr et al. (2008) in 5- to 6-year-olds but not in 10- to 11-year-olds, as the older children were guided by principles of fairness while distributing resources. As a result, older children made conscious efforts to be fair and inclusive but younger children were not able to do the same. Thus, discriminatory attitudes and behaviours might be more easily observable in younger children as compared to older children (Rutland et al., 2005). However, research on prejudice with very young children has also found some inconsistencies. For instance, Howard et al. (2015) found that 3-year olds, but

not younger children, strongly screened out outgroup information and showed such strong dispreference for outgroup individuals that they even avoided toys offered by them.

In sum, a preference for own race has been found even in infants, but has been noted to become more evident in preschool children. These findings suggest that early life (e.g., preschool to the beginning school years) is a time when children have developed clear preference for own race, which may be easily observed in an experimental setting. Therefore, this age might be a particularly good time for assessing prejudice and in testing interventions aiming to reduce it (Gonzalez, et al., 2017).

Implicit and Explicit Bias

Prejudice (a preference for 'own race' over 'other race') can be either implicit or explicit. Implicit prejudice refers to an unconscious, automatic association that affects judgments, yet the person might not be completely aware of them (Baron, 2015) whereas explicit prejudice refers to observable, discriminatory behaviour (McGlothlin & Killen, 2010). There have been mixed results regarding the relation between explicit and implicit bias. For instance, Rutland et al. (2005) reported no significant relation between the two, whereas Newheiser and Olson (2012) found explicit bias to be a significant predictor of implicit bias. A meta-analysis conducted by Hofmann et al. (2005) found a weak relationship between implicit and explicit prejudice. Further, Dovidio et al. (2002) noted that both manifest in different ways. Thus, implicit bias accounts for nonverbal friendly contact and explicit bias accounts for verbal friendliness in an interracial context. It is therefore important to measure both.

Past studies have largely focused on explicit bias measurement and its reduction but recently there has been an interest in studying implicit bias. By the age of 3 years, if not earlier, children display signs of implicit racial bias (Dunham et al., 2013) and this bias has been found to remain stable and resistant to numerous age-related factors. Explicit bias, on the other hand, has been successfully reduced by intergroup contact in 3- to 5-year-old children (McGlothlin & Killen, 2006).

Parent-child prejudice connection

Recent studies have, by and large, indicated a relation between many different types of parental attitudes and children's prejudiced attitudes (e.g., see the meta-analysis by Degner & Dalege, 2013). Regarding implicit bias, Castelli et al. (2009) found that it is more easily transmitted from parents to young children. We therefore included measures of parent's'attitudes in our study.

There are certain measures of social attitudes that correlate with prejudice, chief amongst these, Social Dominance Orientation (SDO) and Right-Wing Authoritarianism (RWA). SDO assesses one's acceptance of inequalities in society, whereas RWA relates more to one's willingness to submit to authorities perceived as established and legitimate (Duckitt & Sibley, 2010). In New Zealand, where the present study was carried out, SDO has been found to relate to low warmth towards ethnic minorities, and RWA to anti-immigration attitudes (Satherley & Sibley, 2016). These findings have been obtained with adults, but recently, Ruffman et al. (2020) found links between maternal SDO and prejudice in children aged 6 to 12 years as well. Additionally, studies with adolescents and their parents have indicated that parental SDO and RWA are specifically and uniquely related to offspring SDO and RWA respectively (Duriez et al., 2008). Therefore, in the present study, we included specific instruments to measure parents' racial prejudice to examine potential links between parental racial attitudes and pre-schoolers' racial prejudice.

Reducing Interracial Prejudice

Childhood attitudes are argued to be more malleable than those in adults, and therefore, it is important to try and shift negative beliefs about the 'other race' at a young age (Aboud et al., 2012; Raabe & Beelmann, 2011). Prior interventions based on contact theory (Allport, 1954) have successfully reduced prejudice in school-age children by encouraging contact between minority and majority ethnicities (see the meta-analyses by Aboud et al., 2012; Ülger, et al., 2018), and so have studies using imagined contact (Birtel et al., 2019). Thus, we examined whether children's attitudes toward other ethnicities might also be changed, although in our case, we used brief picture books which introduce children to 'other-race' characters for four weeks.

Our specific hypotheses were:

1. An own-race preference will be evident in children as young as 3 years on both implicit and explicit measures.
2. Older children (above 5 years) will show less explicit bias than younger children as indicated by previous research, although implicit prejudice would be maintained (see above).
3. There would be a positive relation between parents' racial attitudes and children's racial attitudes.
4. Reading cross-race friendship books will increase familiarity and reduce racial prejudice towards Muslims and Asians amongst participants in the Experimental (Asian Friendship) group.

METHODS

Participants

The sample consisted of 88 children between 3 and 8 years old (M = 5.18, SD = 1.32). Data were collected between August, 2019 and December, 2019. All of the participants were from Dunedin, New Zealand. Seven children were excluded because they did not identify with the European ethnicity. That is, at the beginning of the experiment, each participant was asked to point to the picture of the child that was similar to them (while being presented with a photograph of two children of the same gender and roughly the same age but who were either European or Asian). Seven children pointed towards the Asian child's photograph. Therefore, we excluded these seven cases from all further analyses. Of the remaining children that constituted our final sample, 36 were boys and 45 were girls. The accompanying parent was almost always the mother, with her highest education recorded as a measure of socio-economic status (M = 2.73; SD = .869, where 1 = some high school, 2 = some professional or vocational training, 3 = undergraduate degree and 4 = post-graduate degree).

Based on evidence suggesting that younger children display greater prejudice than older children (Gonzalez,

Steele, & Baron, 2017; Raabe & Beelmann, 2011), we spilt the data into two age groups: Younger (*M*= 4.08; *SD*= .567 and Older (*M*=6.41; *SD*=.704) with roughly equal number of participants in each age group.

Further, for analyses pertaining to the Implicit Racial Bias Test, we excluded children for various reasons (see below), which left 56 children (*M*age = 5.57, *SD* = 1.30). A post hoc power analysis using G*Power3 (Faul et al., 2007) indicated that to test the experimental group differences in the two age groups, with a medium effect size (*f*=.25), and an alpha of .05, a total of 54 participants were required to achieve a power of .95. Thus, for both explicit and implicit measures, we had enough participants to go ahead with the repeated measures analyses.

Measures

Pre/Post Exposure Tasks for Children

Participants' racial attitudes from three tasks were examined pre- and post-media exposure. These measures included the following tasks.

Explicit Race Preference Task. We asked children short questions about who they thought was kind or helpful (Appendix A). For example, one of the stories was about a 'kind' boy who saved a kitten from drowning, with children asked which was the kind boy (with two pictures, one displaying an Asian and the other a European child). All the questions had positive adjectives and the participants were asked to pick one of the two options. These positive adjectives were adapted from the Preschool Racial Attitude Measure II (PRAM II) by Williams, et al (1975) and some of these were included in the friendship books as well. All the pictures in this task were matched for age and attractiveness based on the ratings obtained from 20 postgraduate students. There were five pairs of girls and five pairs of boys. All responses favouring Europeans were coded as '1' and those favouring Asians were coded as '-1'. The maximum score possible over 10 trials was 10 and the minimum was -10, with a positive score indicating that participants demonstrated a greater preference for their 'own race' and a score of 0 indicating no preference for either race.

Explicit Discrimination Task. In four different scenarios, we asked children how far or close they would like to be to Olivia, Khadija, Adam and Yong Chen (see Appendix A, Discrimination Task). For this task, children were presented with pictures of two European children's pictures (Olivia and Adam) and two Asian children's pictures (Khadija and Yong Chen). The names were kept the same as those that would appear in the books during intervention with the expectation that children from the experimental group would show greater acceptance towards Khadija and Yong Chen after reading the books.

For these tasks, children could choose from 1 (very close) to 7 (as far as possible). Olivia and Adam were European children who appeared in the European friendship books whereas Yong Chen and Khadija were Asian children who appeared in the Asian friendship books. This scale was similar to the one constructed and used by Berger et al. (2015) to assess discriminatory tendencies in the Israeli–Palestinian context although our scale had four items rather than a single item to increase sensitivity.

Implicit Racial Bias Test (IRBT). This task, like the other two, was repeated at pre- and post-intervention. Our version of the IRBT followed that used by Qian et al. (2019) for pre-schoolers (which had been adapted from Cvencek et al., 2011). As with other implicit bias tests, the purpose was to measure whether children had a positive association with their own race and a negative association with the other race. As opposed to the traditional Implicit Association Test, the IRBT requires participants to learn only one set of associations at a time. It also uses images instead of words, which makes it more suitable for pre-schoolers (Danziger & Ward, 2010). This format is similar to that used by researchers to explore implicit gender bias (Cvencek et al., 2011), body shape bias (Thomas et al., 2007) and racial bias (Qian et al., 2019). We measured children's levels of pro-European/anti-Asian bias by calculating how quick they were to pair 'thumbs up' and 'thumbs down' icons with European vs. Asian faces (see Figure 1).

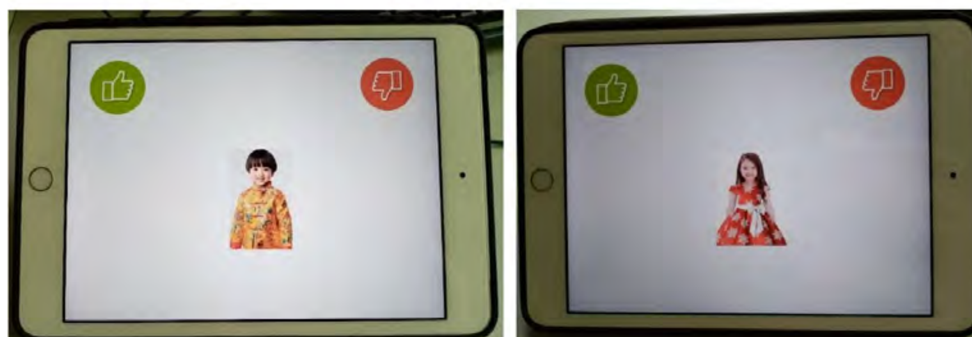


Figure 1. IRBT Apparatus Used by Children

We administered the IRBT using an iPad Mini by Apple Inc., which had a screen size of 7.9 inches and seemed to fit perfectly into tiny hands. The touch screen made it easier for children to select their answers. We used the Perception Research Systems software®. Participants were presented with 20 Asian children (10 boys and 10 girls) and 20 European children (10 boys and 10 girls), with children dressed in culture-consistent dress and headgear in two blocks. These sets of images were matched for age and attractiveness after a pilot test based on the ratings obtained from 20 postgraduate students (details discussed below). Each participant saw one image at a time in the centre of the screen, with options of 'thumbs up' button and a 'thumbs down' button (see Figure 6.1). For congruent pairings, participants were

asked to press the 'thumbs up' button when they saw a child similar to themselves, but the 'thumbs down' icon when they saw someone different. These rules were reversed for the incongruent trial. Approximately half of the participants completed the incongruent trial first while the others completed the congruent trial first to control for any effects of fatigue or practice. We also had one practice block at the beginning of each block to familiarize participants to the format. Therefore, each participant completed a total of four blocks: two practice blocks (eight questions each) and two trial blocks (20 questions each). We replaced each incorrect trial by the mean response time for correct responses and added 600ms penalty following the procedure recommended by Greenwald et al. (2003). In line with prior Implicit Attitude measures, we excluded the practice trials as well as any response latencies above 10,000 ms or below 300 ms. We also excluded any child with an error rate > 60%, and any trial with an average response latency 3 *SD* above the mean response latency (Cvencek et al., 2011; Qian et al., 2019). This left us with 56 participants (see the Participants section).

We computed a *D*-score in accordance with prior researchers (Greenwald et al., 2003; Qian et al., 2019) by using the equation: (RTincongruent-RTcongruent)/*SD*. This *D*-score was used for all further analyses of implicit bias in this study. A *D*-Score of 0 indicates no bias, a positive *D*-Score indicates own-race preference, and a negative *D*-Score indicates other-race preference.

Parental Attitudes

For measuring parental attitudes, we used three measures:

SDO and RWA. SDO was assessed using six items from the original SDO scale (Pratto et al., 1994, see Appendix A). The reason for using the short version was to make sure that parents did not lose motivation with a lengthier version. Other researchers have found this six-item scale to successfully measure SDO (Bergh et al., 2015; Osborne et al., 2017; Stanley et al., 2019). For each of the six questions, parents rated the extent to which they agreed to each statement on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Three items favoured dominance and three favoured equality. After reversing the items that favoured equality, we averaged all six to create one variable for SDO. The SDO scale for the parents had an acceptable internal reliability after deletion of item 3 (α = .745; *M* = 1.83; *SD* = .869)

RWA was assessed using six items from the RWA scale (Altemeyer, 1998, see Appendix A) following other researchers who found this scale to accurately gauge right-wing authoritarian attitudes (Stanley et al., 2019). Parents again rated the extent to which they agreed with each item on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). We then created a composite score after reversing the three reversed items. The RWA scale had an acceptable desirable internal reliability after deletion of item 1 (α =.697; *M* = 2.34; *SD* = 1.10).

Racism, Acceptance and Cultural-Ethnocentrism Scale (RACES). Parent racism was assessed using the RACES (Grigg & Manderson, 2015), a 24-item scale consisting of three subscales (Accepting Attitudes, Racist Attitudes and Ethnocentric Attitudes). RACES was developed bearing in mind the Australian culture so for all questions the country name 'Australia' was replaced with 'New Zealand' in the current study. It consisted questions like, "If people aren't happy with NZ, they should go back to their own country" and "People from all backgrounds are equal". The latter item, along with 11 more items, were reverse-scored, and a composite score was created for all 24 items (α =.843; *M* = 1.75; *SD* = .338). Parent Racism positively correlated with SDO, *r* = .22; *p* =.049 but not with RWA, *r* = .206; *p* =.067.

Procedure

Participants were tested before and after four weeks of reading friendship storybooks outlining friendship between two children who were either all European (which will be referred to as European Friendship group), or between children who were European and Asian/ Arab Muslim (which will be referred to as Asian Friendship group from hereon). In the first session, we measured children's explicit racial attitudes with two measures (see Appendix A) and recorded their implicit racial attitude using the Implicit Race Bias Test (IRBT) that we had created. We also obtained the parent's SDO, RWA and Racism (RACES) scores in the first session (see Appendix B). Children were then randomly assigned two books (either about European Friendship or about Asian friendship depending on which experimental group they had been randomly assigned to. These short picture books had the same text, with only the names and images (race) of the characters differing. The Experimental group (*n* = 45) received stories about friendship with an Arab and Chinese child namely *'How I met Khadija'* and '*How I met Yong Chen'*. The Control group (*n* = 36) received stories about two European children named '*How I met Olivia*' and '*How I met Adam'*. Both stories followed the same pattern of initial anxiety when the new character was introduced and then eventually a long-lasting friendship after learning that they had similar interests and desires. A few positive adjectives from PRAM II by Williams et al (1975) were included in these books. These were the same that would be measured in the explicit racism task prior to and after reading these books.

The parents were instructed to read the books to the children, on average three times a week for four weeks as prior research indicates that reading stories across four to six settings is more effective in reducing bias (Aronson et al., 2015; Cameron et al., 2006). Parents were also given a reading schedule sheet to record the number of times they had read each book and to note any comments made by the child. For this experiment, the average number of times a participant had read the books was, *M* = 13 times; *SD* = 8.12. After this, the participants were retested on the explicit and implicit tasks.

RESULTS

The descriptive statistics for all the measures are presented in Table 1.

Implicit Bias Results

D-Scores scores were normally distributed, as assessed by Shapiro-Wilk's test (*p* > .05) and there were no outliers in the data, as assessed by inspection of a boxplot. We explored IRBT *D*-score in a 2 (Age Group: 3- & 4-year-olds, 5-years +) x 2 (Experimental Group:

European friendship, Asian friendship) x 2 (Time: pre-test, post-test) mixed analysis of variance (ANOVA). Age Group and Experimental Group were between-subjects variables and Time was a within-subjects variable. There was no main effect for Time, F (1, 65) = .125, p = .725, η_p^2 = .002, Experimental Group, F (1, 65) = 1.71, p= .195, η_p^2=.026, or Age Group, F (1, 65) = .520, p = .473, η_p^2 = .008. Neither were there any significant interactions (all Fs < 1.59 and all ps > .212).

Although D-scores were not affected by the Experimental Group or Age, they did indicate bias. Thus, when collapsing over Experimental Group and age, the D-score at Time 1 (M = .246; SD = .849) was significantly different from a no-bias score of 0, t (55) = 2.17, p = .035, Cohen's d = .29.

Explicit Bias Results

Next, we explored Explicit Race Preference in a 2 (Age Group: 3- & 4-year-olds, 5-years +) x 2 (Experimental Group: European friendship, Asian friendship) x 2 (Time: pre-test, post-test) mixed analysis of variance ANOVA. There was no main effect for Time, F (1, 79) = .115, p = .736, η_p^2 = .001, Experimental Group, F (1, 79) = 1.43, p = .707, η_p^2 = .002, or Age Group, F (1, 79) = .012, p = .912, η_p^2 = .000. Neither were there any significant interactions (all Fs < .171 and all ps > .681). Although Race Preference scores were not affected by the Experimental Group or age, they nevertheless indicated bias. Thus, collapsing over Experimental Group, and when examining pre-intervention scores, the Race Preference score (M = 2.28; SD = 3.59) was significantly different from 0 (with 0 indicating no bias and a positive score indicating an own-ethnicity bias), t (78) = 5.64, p < .001, Cohen's d = .63.

Next, we explored Explicit Discrimination in a 2 (Time: pre-test, post-test) x 2 (Age Group: 3- & 4-year-olds, 5-years +) x 2 (Experimental Group: European friendship, Asian friendship) mixed methods ANOVA. Time was a within-subjects variable whereas Age and Experimental Groups were between-subject factors. The results of this analysis indicated that there was no main effect for Time, F (1, 77) = .007, p = .933, η_p^2 < .001, Experimental Group, F (1, 77) = .029, p = .866, η_p^2 < .001, or Age Group, F (1, 77) = .361, p = .550, η_p^2 = .005. Neither were there any significant interactions (all Fs < 2.13 and all ps > .334).

Finally, we used multiple regression analysis to predict the D-score at Time 2, with Time 1 ratings, Parent SDO, Parent RWA and Parent Racism as predictors. All

Table 1. Descriptive Statistics for Explicit and Implicit Bias

	Dependent variable	Asian friendship group *M* (*SD*)		European friendship group *M* (*SD*)	
		Younger	Older	Younger	Older
Time 1	Explicit Race Preference	1.92 (2.80)	3.06 (3.61)	2.63 (3.56)	1.80 (4.54)
	Explicit Discrimination Task	-0.22 (3.03)	1.61 (4.85)	1.50 (3.69)	0.45 (3.63)
	Implicit IRBT D-Score	.167 (.980)	.414 (.847)	.550 (1.62)	.288 (.672)
Time 2	Explicit Race Preference	1.63 (3.76)	2.82 (5.25)	2.63 (2.60)	1.00 (4.27)
	Explicit Discrimination Task	0.48 (4.14)	1.17 (3.07)	0.75 (3.57)	0.75 (3.16)
	Implicit IRBT D-Score	.404 (.867)	.051 (.539)	.722 (1.35)	.357 (.910)

Table 2. Multiple Regression Analyses with Race Discrimination, D-Score and Race Preference at Time 2 as Dependent Variables

	IRBT D-Score			Explicit Race Preference			Discrimination		
	B	*SE*	*β*	*B*	*SE*	*β*	*B*	*SE*	*β*
Time 1 Value	.329	.118	.339**	.517	.120	.459***	.113	.110	.121
Parent Racism	.758	.349	.269*	-.567	1.32	-.047	-1.26	1.28	-.118
Parent SDO	-.154	.143	-.135	.596	.509	.128	.617	.484	.151
Parent RWA	-.117	.122	-.131	-.437	.388	-.121	.026	.383	.008

Note. *p ≤ .05, **p ≤ .01, ***p ≤ .001.

predictors were entered in a single step. Together, these variables predicted a significant amount of variance in the Time 2 D-score, F (4, 62) = 5.38, p = .001), with an R^2 of .258. Time 1 Ratings and Parental Racism were significant predictors on their own, even after accounting for the variance explained by the other variables (p = .006 and p = .011 respectively). We then used two similar analyses to examine Time 2 Race Preference and Time 2 Asian Discrimination. The results are presented in Table 6.2.

DISCUSSION

In this experiment, we examined whether young children in New Zealand displayed a preference for own-race over other-race and if that could be reduced by brief but repetitive media exposure to Asian and Muslim characters via short cross-race friendship picture books. The results indicated that children as young as 3 years old consistently demonstrated a preference for their own race at both an explicit and implicit level. These results are in line with what other researchers have found for children of this age range (Baron & Banaji, 2006; Dunham et al., 2006; Qian et al., 2016) However, this study is unique as no study has examined anti-Asian and anti-Muslim bias in NZ children. The results of our study indicate that children display the same trends in racial and religious prejudice as adults in NZ.

Further, the results also indicated that the preference for own race was hard to change in pre-schoolers such that that exposure to books about friendship with Asian characters (13 times on average) was not enough to reduce prejudice. Children continued to choose own-race (European) playmates and associated positive adjectives with them more readily than with Asian or Muslim children. Moreover, we did not find any effects between the Age Groups for either implicit or explicit bias in this study. Some other researchers have also arrived at similar results (e.g. Dunham et al., 2006; Setoh et al., 2017; Qian et al., 2016) providing evidence that racial biases emerge early and remain stable throughout life. Our results also support that implicit bias may be consistent across early childhood. Further, contrary to our hypothesis that explicit bias would reduce over age, we found that it was not the case. Perhaps 5- and 6-year olds (who constituted the older children in this study) had not yet achieved the level of fairness that Fehr et al. (2008) noted in 10-year-olds which marks the transition towards becoming consciously aware and thus avoiding race discrimination in everyday life or at least attempting to conceal it in an experimental setting.

Parental RWA and SDO were unrelated to child attitudes but parent racism was related to implicit own-race bias in children, emphasising the importance of intergenerational transmission of prejudiced attitudes. This finding is similar to Castelli et al. (2009) who found that children are able to pick up non-verbal cues from their parents. Thus, they may display a bias very similar to their parents, but without being aware of it. Taken together, our results indicate that children as young as 3 years old display anti-Asian and anti- Muslim prejudice that is relatively inflexible, and that parental attitudes are important in shaping children’s racist attitudes at least at an implicit level.

One limitation of this study was that parents were reading the picture books to their young children. Recent studies have addressed the role of communicator in the process of bias reduction. For instance, Endendijk et al. (2014) found that mothers were subtly communicating gender stereotypes to their children while reading storybooks. Further, regarding racial bias, Pahlke et al. (2012) found that while reading books that encourage discussion about interracial friendships, mothers shifted the focus to nonhuman friendships or other goals unrelated to race or discrimination. Indeed, other researchers have also found that White parents specifically refrain from discussing race with their children (Katz, 2003; Vittrup & Holden, 2011). This may have been the reason that children appeared to miss the emphasis on cross-race friendships and therefore reading the books made no difference to their racial attitudes.

Further, during the data collection, we observed that many parents were reluctant to discuss race-related issues with their young children, some even explicitly remarked that it was unnecessary to read cross-race friendship books. Contrary to this view, there is evidence that books introducing children to different cultural narratives are important for developing diversity awareness and assisting in identity development, critical literacy and empathy (Drucker, 2003; Myers, 2014). Future research should therefore continue to explore how children’s attitudes would be affected by such cross-race friendship books when parents are excluded from this process altogether.

It is also noteworthy that children growing up in Dunedin may have lesseexposure to cross-race individuals, perhaps substantially less than children growing up in Auckland, for instance, which has been rated as one of the world's most culturally diverse cities (International Organization for Migration, 2015). Therefore, it is possible that children in other cities, which have a greater in-person interaction with different races, show greater acceptance of other races. There is strong evidence that exposure to diverse cultures in childhood reduces prejudice and this would be an area worth exploring by conducting a similar study in other, more culturally diverse cities.

References

Aboud, F. E., Tredoux, C., Tropp, L. R., Brown, C. S., Niens, U., & Noor, N. M. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review, 32*(4), 307–336. doi:10.1016/j.dr.2012.05.001

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., Tanaka, J. W., & Lee, K. (2013). Developmental Origins of the Other-Race Effect. *Current Directions in Psychological Science, 22*(3), 173–178. doi:10.1177/0963721412474459

Ayling, L. (2020, January 14). Mother claims her daughter was called a 'stinky Indian', had DOG FAECES rubbed in her face by other children and was threatened with a fake gun at a popular park. *Daily Mail*. Retrieved from https://www.dailymail.co.uk

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and Nurture in Own-Race Face Processing.

Psychological Science, 17(2), 159–163. doi:10.1111/j.1467-9280.2006.01679.x

Baron, A. S. (2015). Constraints on the development of implicit intergroup attitudes. *Child Development Perspectives, 9*, 50–54. https://doi.org/10.1111/cdep.12105

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science, 17*, 53–58. https://doi.org/10.1111/j.1467-9280.2005.01664.x

Berger, R., Benatov, J., Abu-Raiya, H., & Tadmor, C. T. (2015). Reducing prejudice and promoting positive intergroup attitudes among elementary-school children in the context of the Israeli-Palestinian conflict. *Journal of School Psychology, 57*, 53–72. https://doi.org/10.1016-j.jsp.2016.04.003

Bian, L., Sloane, S., & Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proceedings of the National Academy of Sciences, 115*(11), 2705–2710. doi:10.1073/pnas.1719445115

Bigler, R. S., & Liben, L. S. (1993). A cognitive-developmental approach to racial stereotyping and reconstructive memory in Euro-American children. *Child Development, 64*(5), 1507–1518. https://doi.org/10.2307/1131549

Birtel, M. D., Di Bernardo, G. A., Stathi, S., Crisp, R. J., Cadamuro, A., & Vezzali, L. (2019). Imagining contact reduces prejudice in preschool children. *Social Development, 28*(4), 1054–1073. doi:10.1111/sode.12374

Burkholder, A. R., D'Esterre, A. P., & Killen, M. (2019). Intergroup Relationships, Context, and Prejudice in Childhood. *Handbook of Children and Prejudice*, 115–130. doi:10.1007/978-3-030-12228-7_6

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective Imitation of In-Group Over Out-Group Members in 14-Month-Old Infants. *Child Development, 84*(2), 422–428. doi:10.1111/j.1467-8624.2012.01860.x

Cameron, L., Rutland, A., Hossain, R., Petley, R. (2011). When and why does extended contact work? The role of high quality direct contact and group norms in the development of positive ethnic intergroup attitudes amongst children. *Group Processes & Intergroup Relations, 14*, 193-206. doi:10.1177/1368430210390535

Cooley, S., Burkholder, A. R., & Killen, M. (2019). Social inclusion and exclusion in same-race and interracial peer encounters. *Developmental Psychology, 55*(11), 2440–2450. https://doi.org/10.1037/dev0000810

Cvencek, D., Greenwald, A. G., & Meltzoff, A. N. (2011). Measuring implicit attitudes of 4-year-olds: The Preschool Implicit Association Test. *Journal of Experimental Child Psychology, 109*(2), 187–200. doi:10.1016/j.jecp.2010.11.002

De Bruijn, Y., Amoureus, C., Emmen, R. A. G., & Mesman, J. (2020). Interethnic Prejudice against Muslims among White Dutch Children. *Journal of Cross-Cultural Psychology, 51*(3-4), 203–221. doi:10.1177/0022022120908346

Degner, J., Dalege, J. (2013). The apple does not fall far from the tree, or does it? A meta-analysis of parent-child similarity in intergroup attitudes. *Psychological Bulletin, 139*(6), 1270–1340. https://doi.org/10.1037/a0031436

Dovidio, J. F., Kawakami, K., Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. Journal of Personality and *Social Psychology, 82*(1), 62–68. https://doi.org/10.1037//0022-3514.82.1.62

Duckitt, J., & Sibley, C. G. (2010). Personality, Ideology, Prejudice, and Politics: A Dual-Process Motivational Model. *Journal of Personality, 78*(6), 1861–1894. doi:10.1111/j.1467-6494.2010.00672.x

Dulin-Keita, A., Hannon, L., Fernandez, J. R., Cockerham, W. C. (2011). The defining moment: Children's conceptualization of race and experiences with racial discrimination. *Ethnic and Racial Studies, 34*, 662–682. doi:10.1080/01419870.2011.535906

Dunham, Y., Baron, A. S., & Banaji, M. R. (2008). The development of implicit intergroup cognition. *Trends in Cognitive Sciences, 12*, 248–253. doi:10.1016/j.tics.2008.04.006

Dunham, Y., Chen, E. E., & Banaji, M. R. (2013). Two Signatures of Implicit Intergroup Attitudes. *Psychological Science, 24*(6), 860–868. doi:10.1177/0956797612463081

Duriez, B., Soenens, B., & Vansteenkiste, M. (2008). The intergenerational transmission of authoritarianism: The mediating role of parental goal promotion. *Journal of Research in Personality, 42*(3), 622–642. doi:10.1016/j.jrp.2007.08.007

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature, 454*(7208), 1079–1083. doi:10.1038/nature07155

Fontanella-Nothom, O. (2019). "Why Do We Have Different Skins Anyway?": Exploring Race in Literature with Preschool Children. *Multicultural Perspectives, 21*(1), 11–18. doi:10.1080/15210960.2019.1572485

Gonzalez, A. M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development, 88,* 123– 130. https://doi.org/10.1111/cdev.12582

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, 15*(2), 260–279. doi:10.1080/1177083x.2020.1733032

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Gregg, A. P., Seibt, B., & Banaji, M. R. (2006). Easier done than undone: Asymmetry in the malleability of implicit preferences. *Journal of Personality and Social Psychology, 90*, 1–20. doi:10.1037/0022-3514.90.1.1

Grigg, K., & Manderson, L. (2015). Developing the Australian Racism, Acceptance, and Cultural-Ethnocentrism Scale (RACES). *The Educational and Developmental Psychologist, 32*(1), 71–87. doi:10.1017/edp.2015.7

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*(10), 1369–1385. https://doi.org/10.1177/0146167205275613

Howard, L. H., Henderson, A. M. E., Carrazza, C., & Woodward, A. L. (2015). Infants' and young children's imitation of linguistic in-group and out-group informants. *Child Development, 86*, 259 –275. http://dx.doi.org/10.1111/cdev.12299

International Organization for Migration (2015). World Migration Report. Retrieved from https://publications.iom.int/system/files/wmr2015_en.pdf

Johnson, P. J., & Aboud, F. E. (2017). Evaluation of an intervention using cross-race friend storybooks to reduce prejudice among majority race young children. *Early Childhood Research Quarterly, 40,* 110–122. doi:10.1016/j.ecresq.2017.02.003

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., … Pascalis, O. (2005). Three-month-olds, but not new-borns, prefer own-race faces. *Developmental Science, 8*(6), F31–F36. doi:10.1111/j.1467-7687.2005.0434a.x

Kinzler, K. D., & Spelke, E. S. (2011). Do infants show social preferences for people differing in race? *Cognition, 119*, 1–9.https://doi.org/10.1016/j.cognition.2010.10.019

Lee, K. J. J., Esposito, G., & Setoh, P. (2018). Preschoolers Favor Their Ingroup When Resources Are Limited. *Frontiers in Psychology, 9.* doi:10.3389/fpsyg.2018.01752

McGlothlin, H., & Killen, M. (2006). Intergroup Attitudes of European American Children Attending Ethnically Homogeneous Schools. *Child Development, 77*(5), 1375–1386. doi:10.1111/j.1467-8624.2006.00941.x

McGlothlin, H., & Killen, M. (2010). How social experience is related to children's intergroup attitudes. *European Journal of Social Psychology, 40*, 625–634. https://doi.org/10.1002/ejsp.733

Newheiser, A.-K., and Olson, K. R. (2012). White and black American children's implicit intergroup bias. *Journal of Experimental Social Psychology, 48*, 264-270. https://doi.org/10.1016/j.jesp.2011.08.011

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*(4), 741–763. doi:10.1037/0022-3514.67.4.741

Qian, M. K., Quinn, P. C., Heyman, G. D., Pascalis, O., Fu, G., & Lee, K. (2017). A Long-Term Effect of Perceptual Individuation Training on Reducing Implicit Racial Bias in Preschool Children. *Child Development, 90*(3). doi:10.1111/cdev.12971

Quinn, P. C., Lee, K., Pascalis, O., & Tanaka, J. W. (2015). Narrowing in categorical responding to other-race face classes by infants. *Developmental Science, 19*(3), 362–371. doi:10.1111/desc.12301

Raabe, T., & Beelmann, A. (2011). Development of Ethnic, Racial, and National Prejudice in Childhood and Adolescence: A Multinational Meta-Analysis of Age Differences. *Child Development, 82*(6), 1715–1737. doi:10.1111/j.1467-8624.2011.01668.x

Renno, M. P., & Shutts, K. (2015). Children's social category-based giving and its correlates: Expectations and preferences. *Developmental Psychology, 51,* 533–543. https://doi.org/10.1037/a0038819

Ruffman, T., Ruffman, C., Hill, S., Turunc, G., Park, N., Du, K., … Hunter, J. A. (2020). RWAc and SDOc: The measurement of right-wing authoritarianism and social dominance orientation in childhood. *Social Development, 29,* 1194-1214. doi: 10.1111/sode.12438 doi:10.1111/sode.12438

Rutland, A., Cameron, L., Milne, A., & McGeorge, P. (2005). Social norms and self-presentation: Children's implicit and explicit intergroup attitudes. *Child Development, 76*, 451–466.https://doi.org/10.1111/j.1467-8624.2005.00856.x

Satherley, N., & Sibley, C. G. (2016). A Dual Process Model of attitudes toward immigration: Predicting intergroup and international relations with China. International *Journal of Intercultural Relations, 53*, 72–82. doi:10.1016/j.ijintrel.2016.05.008

Shutts, K., Kinzler, K.D., McKee, C.B., & Spelke, E.S. (2009) Social information guides infants' selection of foods. *Journal of Cognition and Development, 10*:1-2, 1-17, doi: 10.1080/15248370902966636

Shutts, K., Pemberton, C. K., & Spelke, E. S. (2013). Children's Use of Social Categories in Thinking About People and Social Relationships. *Journal of cognition and development, 14(*1), 35–62. https://doi.org/10.1080/15248372.2011.638686

Tagliaferri, B. (2005). Perception Research Systems Incorporated [Computer software]. Retrieved from www .paradigmexperiments.com.

Thomas, S., Burton Smith, R., Ball, P. (2007). Implicit attitudes in very young children: An adaptation of the IAT. *Current Research in Social Psychology, 13,* 75–85. doi:10.1.1.446.233.

Ülger, Z., Dette-Hagenmeyer, D. E., Reichle, B., & Gaertner, S. L. (2018). Improving outgroup attitudes in schools: A meta-analytic review. *Journal of School Psychology, 67*, 88–103. https://doi.org/10.1016/j.jsp.2017.10.002

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. 48*(1):133.

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2017). Older but not younger infants associate own-race faces with happy music and other-race faces with sad music. *Developmental Science, 21*(2), e12537. doi:10.1111/desc.12537

Corresponding Author
Anita Azeem
Department of Psychology, University of Otago
PO Box 56
Dunedin
Aotearoa/New Zealand
Email: anita.azeem@otago.ac.nz

Appendix A: Explicit Prejudice Tasks

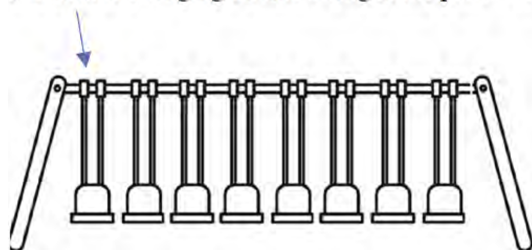
Race Preference Task

1. Here are two boys. One of them is a kind little boy. Once he saw a kitten fall into a lake and he picked up the kitten to save it from drowning. Which is the kind boy?
2. One of these is a smart boy. The other day someone broke the TV remote and he fixed it all by himself. Who is the smart boy?
3. Here are two boys. One of them is a friendly boy. He has a lot of friends. Which one is the friendly boy?
4. Here are two girls. One of them is a helpful girl. She helps her friends and is always there when they need her. Which one is the helpful girl?
5. Here are two girls. One of them is a happy girl. She smiles almost all the time. Which one is the happy girl?
6. One of these is a good girl. She does nice things for her friends and family. Which one is the good girl?
7. Who would you like to go to go camping with?
8. Who would you like to go to the supermarket with?
9. Who would you like to play with at the playground?
10. Who would you like to be friends with?

Child Racism= Composite score across all these questions. Same race responses coded as 1 and different race coded as -1. Responses may range from -10- 10 with a positive score indicating preference for 'own race', a negative score indicating a preference for 'other race' and a score of 0 indicating no racial preference.

Discrimination Task

a. Olivia is swinging on this swing at the park. Where would you like to swing?



b. Khadija lives in this house. Where would you like to live?

c. In a hospital waiting room, you see Yong Chen sitting on this chair. Where would you like to sit?

d. At the library, Adam is sitting on this bean bag in a reading corner. Where would you like to sit?

Appendix B: Measuring Parental Attitudes

Social Dominance Orientation (SDO) and Right- Wing Authoritarianism (RWA) Scales
Instruction: Show how much you favour or oppose each idea below by selecting a number from 1-7 on the scale below. You can work quickly, your first feeling is generally best.

1	2	3	4	5	6	7
Strongly oppose	Somewhat oppose	Slightly oppose	Neutral	Slightly favour	Somewhat favour	Strongly favour

1. An ideal society requires some groups to be on top and others to be on the bottom.
2. Some groups of people are simply inferior to other groups.
3. No one group should dominate in society.
4. Groups at the bottom are just as deserving as groups at the top.
5. Group equality should not be our primary goal.
6. It is unjust to try to make groups equal.
7. We should do what we can to equalize conditions for different groups.
8. We should work to give all groups an equal chance to succeed.
9. It is always better to trust the judgment of the proper authorities in government and religion than to listen to the noisy rabble-rousers in our society who are trying to create doubt in people's minds.
10. It would be best for everyone if the proper authorities censored magazines so that people could not get their hands on trashy and disgusting material.
11. Our country will be destroyed someday if we do not smash the perversions eating away at our moral fibre and traditional beliefs.
12. People should pay less attention to The Bible and other old traditional forms of religious guidance, and instead develop their own personal standards of what is moral and immoral.
13. Atheists and others who have rebelled against established religions are no doubt every bit as good and virtuous as those who attend church regularly.
14. Some of the best people in our country are those who are challenging our government, criticizing religion, and ignoring the "normal way" things are supposed to be done.