# Classifying Ethnicity in Multi-Ethnic Contexts

## Implications of Methodological Decisions on Quantitative Research

**Esther S. W. Yao**

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Education,

The University of Auckland, 2022.

# Abstract

Ethnicity is an important variable in health and social science research, but classifying ethnicity is complex because of the construct's fluidity and multiplicity. There are several possible ethnic classification methods, each with strengths and limitations. Existing literature suggests that researchers' decisions on ethnic classification method can influence research findings. However, there is little empirical research on how these classification methods affect quantitative research in samples with higher multi-ethnic prevalence (e.g., >10%). Situated within a critical quantitative paradigm, the current thesis addresses this gap by utilising large-scale datasets from increasingly multi-ethnic cohorts (adults, adolescents, and children) in Aotearoa New Zealand to examine the effects of ethnic classification method on reported ethnic group size, demographic composition, and substantive outcomes.

The corresponding three studies found that, first, ethnic classification method can substantially impact outputted ethnic group sizes, especially when there are higher rates of multi-ethnic identification (e.g., among children and adolescents). Second, there was a high rate of discrepancy between the two popular methods of administrative-prioritisation and self-prioritisation in each age cohort (≥60%). These discrepancies were systematically associated with contextual characteristics such as neighbourhood ethnic composition and socioeconomic deprivation. Third, using adolescent mental health outcomes as a case study, it was found that ethnic classification method can affect substantive outcomes, both within nominal ethnic groups (by an effect size of up to $d = 0.12$), and between nominal ethnic groups (by an effect size of up to $d = 0.25$).

Together, the results indicate that researchers' choice of ethnic classification method affects who is included or excluded from ethnic groups, the demographic composition of these ethnic groups, and conclusions about the extent of ethnic disparities. This, in turn, influences

knowledge construction, practice, and policy. Given that each ethnic classification method has strengths and limitations, the studies in this thesis highlight the importance for researchers to critically and transparently select the most appropriate ethnic classification method for their research question and context. The power of the researcher, combined with the subjectivity of ethnicity measurement, also emphasises the need for researchers to be critical and reflexive throughout the wider research process.

*Keywords*: race/ethnicity, multiple ethnicities, ethnic measurement, ethnic classification, quantitative methods, critical quantitative research

# Acknowledgements

This thesis, and indeed my PhD journey both academically and personally, would not have been possible without the following people:

First and foremost, to my supervisors Drs Kane Meissel and Pat Bullen—I cannot even begin to imagine where I would be without you. Thank you both so much for seeing something in me from the beginning, for all the opportunities you have given me, for your mentorship throughout the years, and for always believing in me. Among many other things, Kane—thank you also for introducing critical quantitative research to me; and Pat—for inviting me to the Youth2000 seminar at the beginning of my PhD. These significantly shaped the trajectory of this thesis. Words cannot express how thankful I am for you both.

To Associate Professor Polly Atatoa Carr and Professor Terryann (TC) Clark—thank you so much for all your advisory input and support throughout my PhD. My thinking and worldview have developed immensely because of you both. Many thanks also to Associate Professors Jemaima Tiatia, Theresa (Terry) Fleming, and Roshini Peiris-John, and Professor Susan Morton, for your valuable contributions to the journal manuscripts included in this thesis; and to Professors Gavin Brown and Melinda Webber, and Drs Sarah-Jane Paine, Yvonne Ualesi, and Esther Fitzpatrick, for contributing to my thinking and scholarship. To the Growing Up in New Zealand and Youth'12 participants and their whānau, as well as the respective research and operation teams—thank you for making the studies in this thesis possible. I would also like to acknowledge the anonymous reviewers for their feedback on the journal manuscripts.

To the H306 whānau and Research Group members, both past and present—I am incredibly grateful for your companionship on this journey. I have learnt from, and been inspired by, each of you. Special thanks to Drs Shengnan (Serene) Wang and Joohyun (Justine)

Park, and soon-to-be-Dr Georgia Rudd, for your friendship and support. To my mentors and friends outside of academia—you know who you are—a very special shout-out to you too! Thank you for always being there for me, and for reminding me that there is life beyond these ivory towers.

To my family—I could not have done this without your understanding and support. Dad, thank you for always being eager to read my first drafts and to chat about ideas. Mum, thank you for your listening ear, humour, unconditional support, and encouraging me to celebrate small milestones. Stephanie, of course I have to mention your delicious cooking—brain food! Thank you also for gaming with me when I needed a break from things. Keniel, thank you for getting me into R, and for the fun R bootcamps we had during your summers back home.

Last but not least, to my Heavenly Father, who has blessed me with all these, sustained me through it all, and given me a secure identity independent of socially constructed categories—I am forever grateful.

# Contents

# List of Tables

# List of Figures

# List of Acronyms

| | |
|---|---|
| ANOVA | Analysis of variance |
| CANZUS | Canada, Australia, New Zealand, and the United States |
| CI | Confidence interval |
| GLM | General linear model |
| MEIM | Multigroup Ethnic Identity Measure |
| MGCFA | Multigroup confirmatory factor analysis |
| MELAA | Middle Eastern, Latin American, and African |
| NEI | Not elsewhere included |
| NHIS | (U.S.) National Health Interview Survey |
| NZDep | New Zealand Deprivation Index |
| OECD | Organisation for Economic Co-operation and Development |
| OMB | Office of Management and Budget |
| OR | Odds ratio |
| SDU | Standard deviation units |
| SDQ | Strengths and Difficulties Questionnaire |
| SEM | Structural equation modelling |

CHAPTER 1

# Introduction

As a 1.5-generation Asian-New Zealander who has journeyed with depression since early adolescence, ethnocultural differences in adolescent depression and its implications for practice is an area of special interest to me, and naturally the topic I wanted to research for my PhD. The Youth2000 National Youth Health and Wellbeing Survey Series was an ideal dataset for this purpose, so near the beginning of my PhD, I attended a Youth2000 seminar where there was a presentation on a recent study in adolescent mental health. When outlining the study's method, the presenter mentioned in passing that there are several ways to classify ethnicity, and that they used participants' self-prioritised ethnicity (i.e., their self-reported "main" ethnic group). As someone who had largely experienced ethnicity as a singular and fixed construct, it was my first encounter with the notion of different ethnic classification methods; as an early career researcher fascinated with research methods, I was enthralled. Soon, I found myself fixated on two questions: (1) how will I choose what ethnic classification method to use in my research, and (2) will the method I choose influence my results? These two questions are the antecedent to this thesis.

I searched the literature for answers and found some helpful articles. However, I was not fully satisfied, because the literature which discussed ethnic classification effects were typically either largely theoretical (e.g., Cormack & Robson, 2010; Denton & Deane, 2010; Office of Management and Budget [OMB], 2000), or utilised datasets with relatively low multi-ethnic prevalence (e.g., <10%; Didham & Callister, 2012; Liebler & Halpern-Manners, 2008;

Mays et al., 2003; Ministry of Health, 2008; Te Rōpū Rangahau Hauora a Eru Pōmare, 2000).[1] I wanted to know the empirical impact that my choice of ethnic classification method would have on analyses utilising the Youth'12 data, where over 30% of participants identified with more than one ethnic group (Clark et al., 2013). Therefore, I conducted some preliminary analyses with the Youth'12 dataset to investigate possible effects of different ethnic classification methods on results. At the time, I thought that the best ethnic classification method would become self-apparent, and that I could then use that method to continue my proposed research on adolescent depression. Instead, the more I explored, the more I realised how complex yet important the issue of ethnic classification is, so much so that I subsequently shifted the focus of my PhD research from investigating ethnic differences in adolescent depression, to examining the effects of ethnic classification on applied quantitative research in increasingly multi-ethnic contexts.

Using large-scale survey data from Aotearoa New Zealand with children, adolescents, and adults, the overarching aim of this thesis is to empirically investigate the effects that ethnic classification method has on quantitative research in increasingly multi-ethnic contexts. This is intended to inform quantitative researchers working with ethnicity data of the implications that their decisions can have on analysis, results, interpretations, and conclusions. The following sections in this introductory chapter provide a high-level overview of the thesis by contextualising ethnicity as a variable, summarising the main issues when working with ethnicity data in multi-ethnic contexts, identifying the gaps this thesis seeks to address, and outlining how these gaps will be addressed. The chapter concludes with an overview of the structure of this thesis.

---

[1] For brevity, "multi-ethnic" refers to identification with two or more ethnic groups in this thesis. "Dual-ethnic" will be used when specifically referring to two ethnic groups.

## 1.1 Ethnicity in Context

Ethnicity is a social construct that evolved from a history of racial ideology, where people groups were ordered by "superiority" as defined by the dominant group (Hirschman, 2004; Moya & Markus, 2010; Smedley & Smedley, 2005). As a result of this history and ongoing racism, there are pervasive ethnic inequities affecting Indigenous and ethnic minority groups around the world,[2] including in the health, social, economic, and educational domains (Balestra & Fleischer, 2018; Bécares et al., 2013; Espiner et al., 2021; Gravlee, 2009; Krieger, 2014). These patterns are concerning from an equity and social justice perspective, and need to be addressed. When used appropriately for these purposes, ethnicity is an important variable in research and policy because it can render ethnic inequities visible; improve understanding of the processes underlying these inequities; be used to target resources, interventions, and policies toward positive change; and monitor progress over time (Atatoa Carr et al., 2017; Balestra & Fleischer, 2018; Mays et al., 2003).

Given the importance of ethnicity as a variable, valid and reliable measurement of ethnicity is paramount. However, the construct's multifaceted, fluid, and subjective nature makes it complex to measure (Aspinall, 2018b; Morning, 2008; Roth, 2016). In addition, international migration, interethnic unions, changing personal and societal conceptions of ethnicity, and developments in how ethnicity is measured, are contributing to a steady increase in the number of individuals who self-identify with more than one ethnic group (Aspinall, 2018b; Perez & Hirschman, 2009; Rocha & Aspinall, 2020). For example, in the United States, since the 2000 Census which first allowed respondents to identify with multiple races, the

[2] Two notes on terminology: (1) "inequities" refer to disparities due to unfair or unjust practices, whereas "inequalities" simply refer to numerical differences; (2) "ethnic minority" as used in this thesis encapsulates two aspects—smaller group numbers (i.e., numerical minority), and marginalisation due to oppressive power structures (i.e., the active process of minoritisation).

proportion of the U.S. population which identified with more than one race increased from 2.4% in 2000, to 2.9% in 2010, to 10.2% in 2020 (approximately 33.8 million persons; N. Jones et al., 2021; N. A. Jones & Bullock, 2012).[3] Similarly, in Aotearoa New Zealand, census data show a steadily increasing proportion of individuals who identify with more than one broad ethnic grouping, from 9.0% in 2001, to 10.4% in 2006, 11.2% in 2008, and 11.4% in 2018 (or nearly 540,000 persons; Statistics New Zealand, 2014b, 2020a). The prevalence of multi-ethnic identification is typically higher in younger age groups, and almost all Western societies project a continual increase in multi-ethnic identification rates in their overall population (Aspinall, 2018b; Atatoa Carr et al., 2017; Statistics New Zealand, 2014b). The steadily increasing proportion of multi-ethnic individuals adds additional complexities to the operationalisation of ethnicity for research.

## 1.2 Working with Ethnicity Data

Multiple ethnic identifications add complexity to quantitative research because many common statistical techniques (e.g., chi-square test of independence, analysis of variance [ANOVA], and regression) typically require mutually exclusive categories (Cohen et al., 2003; Field et al., 2012). It is important to emphasise that multi-ethnic individuals should not be interpreted as the "problem", and the onus should be on researchers to respond appropriately to changing demographic patterns. The literature, primarily from the United States (e.g., Herman, 2011; Mays et al., 2003; OMB, 2000) and Aotearoa New Zealand (e.g., Callister et al., 2007; Cormack & Robson, 2010; Statistics New Zealand, 2004), describes several possible methods for researchers to classify multiple ethnic responses for quantitative research. These

[3] The U.S. Census Bureau notes that the substantial increase in multiracial identifications in the 2020 Census is partly due to improvements in how race/ethnicity data were collected and processed (N. Jones et al., 2021).

can broadly be grouped into methods that *retain* multiple ethnic identifications (e.g., total response grouping, where respondents are counted in *each* of the ethnic groups they report, such that the total ethnic count can exceed the total number of respondents; and sole/combination grouping, which respondents are counted *once* according to the group or combination of groups they report); and methods that *reduce* multiple ethnic identifications (e.g., administration-prioritisation, where respondents are assigned to a *single* group based on a *predetermined hierarchy*; and self-prioritisation, where respondents are asked to select their *"main"* ethnic group). According to the literature that discusses the strengths and limitations of different ethnic classification methods, retention methods generally respect participants' multiple ethnic identifications better, whereas reduction methods are generally easier to implement in statistical analysis (Didham, 2005; Denton & Deane, 2010; Subramanian, 2009). Importantly, each ethnic classification method has specific strengths and limitations, so there is no single "best" method (Atatoa Carr et al., 2017; Moubarac, 2013; OMB, 2000; Woo et al., 2011).

There is also a small body of empirical research, again predominantly from the United States and Aotearoa New Zealand, that shows ethnic classification method can influence the size of outputted ethnic groups (Callister et al., 2007; Grieco, 2002; Parker & Makuc, 2002), as well as reported ethnic disparities in health and social outcomes (Boven et al., 2020; Callister et al., 2007; Mays et al., 2003; Rutkowski et al., 2017). These results suggest that researchers' choice of ethnic classification method can impact the interpretations and conclusions drawn from quantitative research. However, existing studies have typically been conducted using samples with relatively low multi-ethnic identification rates (e.g., <10%). Therefore, despite the increasing global rate of multi-ethnic identification, little is known about the empirical impact that different ethnic classification methods have on statistical analysis in samples with higher multi-ethnic rates (e.g., >10%). In addition, despite the popularity of self-prioritisation

in some research disciplines (e.g., education and psychology), previous research has typically only compared the differential effects of total response, sole/combination grouping, and administrative-prioritisation. The differential effects of self-prioritisation on research findings remain largely unknown.

Furthermore, there appears to be a disconnect between the existing literature and mainstream research practice. Despite the effects that ethnic classification method can have on quantitative results, many published studies do not explicitly state the classification method used, suggesting that a lot of researchers are either unaware of the issues related to ethnic classification, or are dismissive of these issues (Bokor-Billmann et al., 2020; Moore, 2020; Moubarac, 2013; Nishina & Witkow, 2020). There are two possible underlying reasons. First, some researchers who do not have lived multi-ethnic experience may be personally unaware of the complexities regarding ethnic classification of multiple ethnic identifications, and academically have not encountered critical discussion in this area. Second, it is possible that some researchers who are aware of the complexities of ethnic classification of multiple ethnicities may dismiss these issues, because they believe there is insufficient empirical evidence that their choice of ethnic classification method will meaningfully impact their research. In both cases, this results in a lack of critical consideration of ethnic classification method. In addition, the ethnic classification method used is unlikely to be stated in research publications, contributing to a self-perpetuating cycle of unawareness regarding the importance of ethnic classification decisions.

## 1.3 Thesis Aims

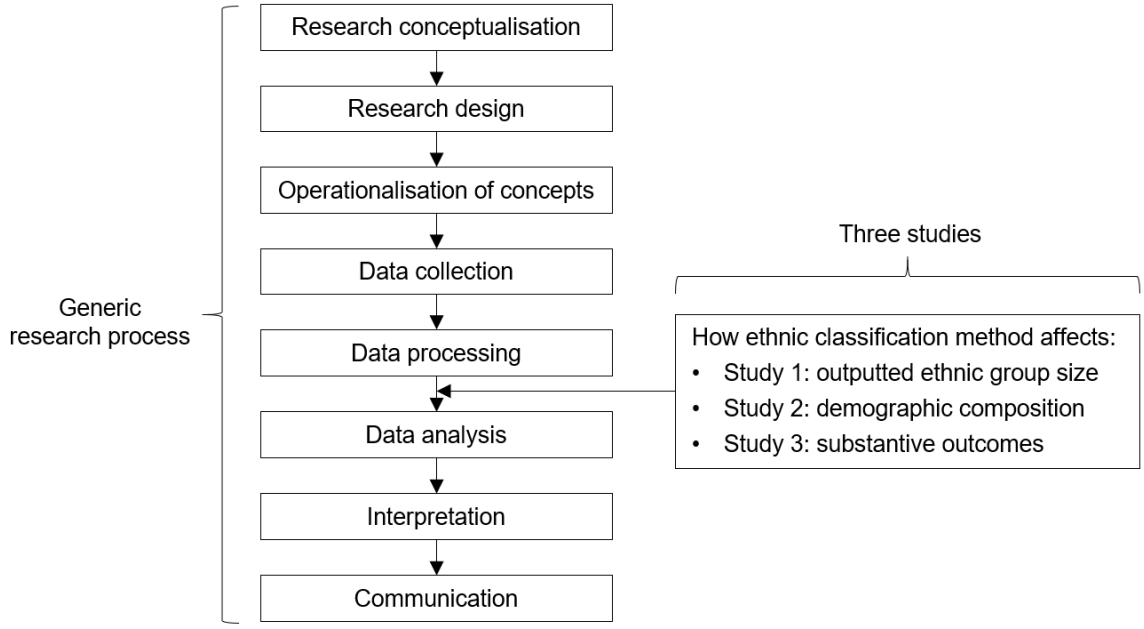
Ethnicity is a crucial variable for equity research and policy, but there is a lack of empirical studies on the effects that ethnic classification method has on the analysis of increasingly multi-ethnic samples, the conclusions drawn from these analyses, and its

implications on research and practice. Using large-scale survey data of children (38% multi-ethnic), adolescents (32% multi-ethnic), and adults (16% multi-ethnic) from Aotearoa New Zealand, this thesis aims to empirically investigate the effects that researchers' ethnic classification decisions have on quantitative analysis in multi-ethnic contexts in three main areas: outputted ethnic group size, demographic composition, and substantive outcomes.

A schematic representation of the studies in this thesis is provided in Figure 1.1. The figure shows a generic research process on the left-hand side, and an overview of this thesis's studies on the right-hand side. Ethnic classification, which is the methodological issue under investigation in this thesis, formally occurs in the data processing stage of the research process—after research conceptualisation, research design, operationalisation of concepts, and data collection; and before data analysis, interpretation, and communication of results. Thus, on one level, this thesis examines how decisions made at the data processing phase affect

**Figure 1.1**

*Schematic Representation of this Thesis's Studies*




*Note.* Research process adapted from Punch (2014).

analyses, and hence, interpretation and communication (e.g., research dissemination). However, the approach to ethnic classification ideally needs to be considered from the outset of the research. Therefore, the thesis also addresses the implications of ethnic classification method on the wider research process, including on research conceptualisation, design, operationalisation of concepts, and data collection.

As shown in Figure 1.1, the effects of ethnic classification method are examined in three areas, culminating in three studies:

- *Study 1*, which examines how ethnic classification method affects *outputted ethnic group size* in increasingly multi-ethnic age cohorts (adults, adolescents, and children)—ethnic group size is important because it represents who is included or excluded in each ethnic group, constitutes the denominator in analyses of outcomes by ethnicity, and is associated with resourcing and funding decisions;
- *Study 2*, which examines how ethnic classification method affects the *demographic composition* of ethnic groups (e.g., age, sex, birthplace, urbanicity, socioeconomic deprivation, and neighbourhood ethnic composition) in each of the three age cohorts—this can reveal whether ethnic classification method can lead to biased samples, and hence biased results; and
- *Study 3*, which examines how ethnic classification method affects *substantive outcomes* in the analysis of ethnic differences in adolescent mental health—that is, whether different ethnic classification methods can result in different results and implications for policy and practice.

I undertake this thesis from the position of an Asian tangata Tiriti (people of *Te Tiriti o Waitangi*, the founding document of Aotearoa New Zealand), and as an ally to Indigenous self-determination and research agenda (Kukutai & Walter, 2015; Smith, 2021). This thesis is situated within a critical quantitative paradigm, which emphasises the importance of examining

quantitative research practices and their underlying influences, with the goals of equity and social justice (Rios-Aguilar, 2014; Stage, 2007; Stage & Wells, 2014).

The purpose of this thesis is threefold:

1. To *empirically investigate the effects of researchers' choice of ethnic classification method* on quantitative analysis, results, and conclusions in increasingly multi-ethnic contexts (i.e., address the gap in the literature);
2. To use these empirical results to *advocate for researcher criticality when selecting an ethnic classification method* (i.e., address the disconnect between the literature and mainstream practice); and more broadly,
3. To *challenge researchers to take a more critical approach when engaging in quantitative research* involving ethnicity data, especially by placing ethnicity in context, being reflexive about how their positionality influences their research, and striving for social justice.

More succinctly, the purpose of this thesis is to critically examine ethnic classification methods, raise consciousness of these issues, and contribute to social change in mainstream quantitative research practice. This thesis will particularly be applicable to researchers who, like me, were trained under a (post-)positivist research paradigm.

## 1.4 Thesis Overview

The structure of the thesis after this introductory chapter (Chapter 1) is as follows. Chapter 2 provides a literature review on ethnicity as a variable, with a focus on general issues related to the thesis. This includes the conceptualisation of ethnicity, the importance of ethnicity as a variable in research and policy, and pertinent concerns regarding ethnic classification internationally and in Aotearoa New Zealand. Literature specific to the individual studies will be presented in their respective chapters (Chapters 4–6). Chapter 3 describes the methodology for the current research, including my researcher positionality, the research paradigm this thesis

is situated within, the datasets used in the empirical studies, and the key ethical considerations related to the research.

Chapters 4 to 6 include the peer-reviewed journal manuscripts for the three empirical studies conducted in this PhD. Chapter 4 (Study 1) examines response patterns to the self-prioritisation question in children (via mother proxy), adolescents, and adults; as well as the effects that different ethnic classification methods have on outputted ethnic group sizes in each age cohort. Chapter 5 (Study 2) delves into the method of self-prioritisation further by exploring the individual and contextual demographic characteristics associated with discrepancies between administrative-prioritisation and self-prioritisation in each age cohort. Chapter 6 (Study 3) focuses on investigating the implications of ethnic classification method on substantive outcomes in adolescent mental health. These three chapters each begin with a preamble, followed by the manuscript in its entirety, with minor edits for consistency. As the manuscripts were written as standalone articles, there is naturally a degree of repetition in the text, particularly in some subsections of the literature review and method. However, I have endeavoured to reduce the overlap to a minimum throughout the thesis. Note that, to emphasise the active role of the researcher, there is occasional use of personal pronouns in the journal manuscripts. These pronouns are in first-person plural form ("we") rather than first-person singular form ("I") due to the manuscripts' co-authored nature.

In Chapter 7, the three individual studies are integrated in a general discussion that considers the overall implications the results have on researchers' choice of ethnic classification method, as well as on working with ethnicity data throughout the research process more broadly. Implications for research audiences are also identified, and the overall limitations and contributions of this thesis are discussed. This is followed by a concluding chapter in Chapter 8, which summarises the thesis and its implications as a whole. A reflection on my academic and personal journey in engaging with ethnicity research is enclosed in the Afterword.

CHAPTER 2

# Literature Review: Ethnicity Matters

The introductory chapter provided a high-level overview of this thesis by contextualising ethnicity, identifying the research gaps in the operationalisation of ethnicity as a variable, and outlining how this thesis will address these gaps. This chapter discusses each of these aspects in more detail. First, from an international perspective, I define ethnicity and related concepts (e.g., ancestry, culture, and race), argue for the importance of ethnicity as a variable with reference to the history of race and persisting contemporary ethnic inequities, review how ethnicity is used in quantitative research, and discuss outstanding issues that need to be addressed. Then, I focus on this thesis's context of Aotearoa New Zealand by outlining its sociohistorical background, summarising its official statistical standard for ethnicity, and discussing the main issues in ethnic classification specific to research in this country. The chapter concludes with a more in-depth outline of the three studies in this thesis. Note the current chapter reviews the literature relevant to this thesis in general; literature specific to the individual studies will be presented in their respective chapters.

## 2.1 Defining Ethnicity and Related Concepts

To begin, it is paramount to establish definitional clarity of the concept of ethnicity and related terms, because conceptual ambiguity and inconsistent definitions have resulted in frequent conflation of these terms (Bhopal, 2004; Cokley, 2007; Morning, 2008). *Ethnicity*—the concept of interest in this thesis—is a social construct that characterises a group of people with commonalities in geographic origin, ancestry, history, traditions, and culture (Bhopal, 2004; Cokley, 2007; Moya & Markus, 2010; Smedley & Smedley, 2005). Although some

scholars note that ethnicity can be socially ascribed (Moya & Markus, 2010; Roth, 2016; White et al., 2020), group membership is increasingly based on self-identification with at least some of these commonalities (Aspinall, 2018b; Balestra & Fleischer, 2018; Morning, 2008; United Nations, 2008). Individuals' ethnic identification(s) are usually collected via a check-all-that-apply question with predetermined ethnic categories (Aspinall, 2018b; Morning, 2008; Rocha & Aspinall, 2020; Roth, 2016). Due to its multi-dimensional and self-perceived nature, ethnic identification is fluid rather than fixed, and can change across time and contexts, especially if an individual identifies with multiple ethnicities (Carter et al., 2009; Didham, 2016; D. R. Harris & Sim, 2002; Liebler et al., 2017; Nishina et al., 2010). For example, an individual's ethnic identification can change depending on their conceptual understanding of ethnicity, their knowledge about their ancestry and culture, their sense of belongingness towards their ethnic group, and the social group they are with at the time (e.g., family vs. peers; Callister et al., 2007; D. R. Harris & Sim, 2002). Note that ethnic identification differs from ethnic identity—ethnic identity refers to individuals' subjective sense of ethnic group membership, and is measured via more detailed questions on their subjective sense of knowledge, belonging, and involvement in their ethnic group(s) (Cokley, 2007; Phinney & Ong, 2007; Roth, 2016).

The concepts of ancestry and culture, which are included in the definition of ethnicity, naturally have some overlap with ethnicity, but each also has fundamental differences. Specifically, *ancestry* focuses on an individual's beliefs about their ancestral descent, whereas ethnicity focuses on an individual's beliefs about themselves (Morning, 2008; Roth, 2016). While ancestry may contribute towards one's ethnic identification, it does not necessitate it. Large-scale studies in the United States and United Kingdom suggest that only around one in three of those who report multiple ancestries also identify with multiple ethnicities (Gullickson & Morning, 2011; Nandi & Platt, 2012; Xu et al., 2021). Similarly, in Aotearoa New Zealand, the 2013 Census showed that around one in six of those who reported Indigenous Māori descent

did not report Māori ethnicity (Statistics New Zealand, 2013). Therefore, ancestry and ethnicity measure interrelated but separate constructs. Conversely, *culture* refers to the ways of life of a group of people, including their language, beliefs, values, customs, dress, and diet (Cokley, 2007; Smedley & Smedley, 2005). Unlike ethnicity and ancestry, culture can be transmitted and acquired via socialisation (Smedley & Smedley, 2005). As a result, individuals can adopt an ethnic group's cultural practices without identifying with the ethnic group itself. Despite differences in these concepts, a key shared characteristic of ethnicity, ancestry, and culture is that, by definition, there is no hierarchy associated with group membership—that is, groups are organised horizontally rather than vertically (Hirschman, 2004; Moya & Markus, 2010; Smedley & Smedley, 2005).

*Race*, on the other hand, is a concept that is inherently intertwined with notions of superiority and inferiority (Hirschman, 2004; Moya & Markus, 2010; Smedley & Smedley, 2005). Traditionally, race referred to a biological construct that hierarchically organises humans into discrete categories based on physical markers, such as skin colour, hair texture, eye shape, and nose shape (Bhopal, 2004; Cokley, 2007; Moya & Markus, 2010; Smedley & Smedley, 2005). According to its original conceptualisation, race is an innate and immutable trait that is responsible for a group's physical appearance, as well as their ability, behaviour, and culture (Smedley & Smedley, 2005). Modern science has largely disproved the biological basis for race (see Section 2.2). Since then, usage of this term in the United States has evolved to incorporate social connotations, leading to conceptual ambiguity (Bhopal, 2004; Cokley, 2007; Morning, 2008). In most other English-speaking countries, the term "ethnicity" (social construct) has superseded "race" (biological construct; Morning, 2008). Accordingly, the term "ethnicity" will be used in this thesis to refer to the social construct. However, "race" will be used when U.S. research is directly cited—in these cases, the word will contain social connotations.

## 2.2 Importance of Ethnicity Data

The contemporary social, research, and policy significance of ethnicity must be understood in the sociohistorical context of race, from which the concept of ethnicity originated. Usage of the word "race" to classify humans began in the late 16th century and was widely used by the 18th century (Hirschman, 2004; Smedley & Smedley, 2005). This practice, and its associated ideology of hierarchy, was developed in the context of increasing Western power and domination, including enslavement of Africans in the Americas; and European colonisation of Asian, African, and Oceanian countries. Racial ideology, which relegated those who were enslaved and colonised to a subordinate status, was used to justify these actions (Hirschman, 2004; Omi & Winant, 2015; Smedley & Smedley, 2005). Around the same time, there was growing Western scientific interest in the taxonomic classification of flora and fauna, which extended to classifying humans into discrete racial groups based on physical characteristics and geography (Hirschman, 2004; Moya & Markus, 2010; Smedley & Smedley, 2005; Tishkoff & Kidd, 2004). Racial "science" (now considered a pseudoscience) employed techniques such as anthropometry, craniometry, and later, intelligence testing, to investigate differences between these racial groups. The findings, which were based on problematic underlying assumptions and failed to account for environmental confounds, "confirmed" the innate inferiority of some groups, adding "validity" to the political agenda of racialisation (Hirschman, 2004; Moya & Markus, 2010; Omi & Winant, 2015; Smedley & Smedley, 2005).

Discourse around race began to shift in the second half of the 20th century, following the Holocaust in World War II, subsequent civil rights and anticolonial movements, and major scientific advancements. In particular, the Human Genome Project found that humans are 99.9% genetically similar, and that there is more variation within races than between races, thus debunking the scientific evidence for discrete racial categories (Jorde & Wooding, 2004; Smedley & Smedley, 2005; Tishkoff & Kidd, 2004). Due to a lack of valid and reliable

biological basis for race, as well as the aforementioned human rights movements, race evolved into the social construct now commonly referred to as ethnicity (Hirschman, 2004; Smedley & Smedley, 2005).

Although race is no longer scientifically accepted as a valid biological concept, its continual legacy and power in the 21st century is reflected in widespread ethnic disparities between dominant and marginalised groups in social, economic, and health outcomes (Balestra & Fleischer, 2018; Bécares et al., 2013; Espiner et al., 2021; Gravlee, 2009; Krieger, 2014). On the surface, these well-documented disparities paradoxically suggest there is a causal effect of race—that is, it appears like these disparities are a result of inherent group differences. However, more nuanced analyses show that the disparities are in fact an indication of the sociocultural reality of persisting racism (albeit usually in more subtle forms; Bonilla-Silva & Zuberi, 2008; Gillborn et al., 2018). Racism, which is based on racial ideology of superiority versus inferiority, can be described at three levels (C. P. Jones, 2000; D. R. Williams, 2018):

- *Institutionalised racism* (also called structural or systemic racism), which occurs when racial groups have differential access to power, opportunities, resources, and services;
- *Interpersonal racism*, which is manifested through stereotypes (prejudice) and differential treatment (discrimination) towards individuals based on their race; and
- *Internalised racism*, where individuals of a marginalised race accept societal messages about their own inferiority.

As an example of the negative consequences of racism, research has found self-reported racial discrimination to be associated with poorer cardiovascular, endocrine, immune, and mental health outcomes (Crengle et al., 2012; Gravlee, 2009; R. Harris et al., 2012; Krieger, 2014). In addition, institutionalised racism contributes to health and social disparities as it negatively impacts individuals' education, employment status, income level, housing quality, healthcare

access, and so on, leading to cycles of intergenerational disadvantage (Gravlee, 2009; Krieger, 2014).

Therefore, due to widespread ethnic inequities, ethnicity is an important and meaningful variable in research and policy. Specifically, ethnicity data are crucial to identifying and quantifying ethnic inequities, understanding the processes underlying these inequities, informing equitable distribution of resources, targeting interventions and policies for equity and justice, and monitoring progress in these areas (Atatoa Carr et al., 2017; Balestra & Fleischer, 2018; Callister et al., 2007; Mays et al., 2003). Morning's (2008) cross-national analysis of 141 countries' ethnic enumeration practices in the 2000 census round showed that over 60% of countries collect some form of ethnicity data, with this being especially prevalent in the Americas and Oceania (>80%). As Rallu et al. (2006) observed, many of these countries previously engaged in racial enumeration to dominate and exclude specific groups, but have since shifted to conducting ethnic enumeration for antiracist and antidiscrimination reasons (e.g., Canada, Australia, Aotearoa New Zealand, and the United States; often collectively referred to as the "CANZUS" colonial settler states). In contrast, most European countries (apart from notable exceptions such as the United Kingdom and Ireland) regard ethnicity as sensitive data (particularly post-World War II), and instead collect nationality or citizenship information for the purposes of national unity (Balestra & Fleischer, 2018; Morning, 2008; Rallu et al., 2006; Simon, 2012). Indeed, there are proponents who argue that "colour-evasive"[4] statistics, where ethnicity is not considered, is a superior method for antidiscrimination than analysing ethnicity data, because ethnicity statistics have traditionally been used to stratify and discriminate (Kertzer & Arel, 2002; Rallu et al., 2006; Simon, 2012). However, the problem

[4] I use the term "colour-evasive" rather than "colour-blind", as the latter—while more common in the existing literature—contains ableist notions (i.e., it associates blindness with [wilful] ignorance; Annamma et al., 2017).

with colour-evasive policies and practices is that it does not eradicate ethnic inequities—rather, it renders it invisible, and therefore, ignorable. Thus, whilst there is evidence that racism and ethnic inequities exist, it is important to measure ethnicity in order to monitor and address it.

On the other hand, there are scholars who contend that it is insufficient to only concentrate on ethnicity if the end-goal is genuine social justice and equity. For example, according to intersectionality theory, it is argued that the focus needs to shift beyond singular constructs to how various social identities (e.g., ethnicity, gender, social class, and sexuality) interact in multiplicative ways in producing privilege or marginalisation (Bowleg, 2012; Crenshaw, 1989; Hancock, 2007). Similarly, some argue that given the "super-diverse" nature of current society, a multi-dimensional perspective of diversity (including ethnicity, country of origin, language, religion, immigration status, and regional identities) needs to be adopted (Boccagni, 2015; Meissner, 2015; Vertovec, 2007). While these arguments are valid and pertinent, ethnicity is clearly still an important variable in these critiques; hence, the classification of ethnicity remains a central underlying issue.

## 2.3 Ethnicity in Research

### 2.3.1 State of the Field

Having established the importance of ethnicity data, the next step is to review how ethnicity has been used in research. An examination of systematic reviews conducted on this topic shows that around two-thirds of published empirical studies which utilise ethnicity data adopt a quantitative approach, usually involving comparison of at least two groups, indicating the particular relevance of exploring how ethnicity is operationalised in quantitative research (Charmaraman et al., 2014; Moubarac, 2013). Concerningly, researchers often did not offer a justification for why ethnic differences were being examined, or provide an explanation for observed ethnic differences. This risks ethnicity being misinterpreted as the cause of observed

differences (i.e., biological essentialism; Bonilla-Silva & Zuberi, 2008), rather than a variable associated with these differences. In articles where an explanation was provided, these varied among socioeconomic, behavioural, environmental, and biological reasons (C. Lee, 2009; Moubarac, 2013).

However, by far the most prominent theme in the identified systematic reviews conducted over the past two decades is the lack of transparency in ethnicity reporting in published articles (Bokor-Billmann et al., 2020; Comstock, 2004; Craven et al., 2018; C. Lee, 2009; Maduka et al., 2021; Møllersen & Holte, 2008; Moore, 2020; Moubarac, 2013). For example, a comprehensive systematic review by Moubarac (2013) on persisting problems in research related to race and ethnicity, which examined 280 articles published in high impact public health and epidemiology journals between 2009 and 2011, found that 64% of the studies examined did not differentiate between the concepts of race and ethnicity. This is problematic because ambiguous and subjective concepts such as race and ethnicity must be clearly defined in order for them to be accurately measured and appropriately interpreted. Regarding ethnicity measurement, Moubarac also found that 42% of studies did not indicate how participants were classified to the racial/ethnic categories used, which he posits implies many researchers are unaware of the complexities associated with ethnic classification, such as the fluidity and multiplicity of ethnic identifications. Based on these findings, Moubarac emphasises that "authors must be transparent in their methods to ascribe race or ethnicity and acknowledge the limits of the classification they choose to use" (p. 113). While the studies examined in Moubarac's comprehensive review are now over a decade old, more recent systematic reviews, which have typically focused on the use of race/ethnicity in narrower fields such as gastroesophageal reflux disease (Craven et al., 2018), ophthalmology (Moore, 2020), medicine (Bokor-Billman et al., 2020), and surgery (Maduka et al., 2021), have found similar results regarding the lack of transparency in reporting.

To help address the limitations identified in these systematic reviews, there is ample literature to guide researchers on conceptual issues around race and ethnicity (see Section 2.1 and references there), and reasonable attention directed towards how researchers should collect ethnicity data in increasingly multi-ethnic contexts. For example, the United Nations' (2008) *Principles and Recommendations for Population and Housing Censuses* states that, due to the subjective nature of ethnicity and increasing multi-ethnic prevalence, ethnicity data should be obtained via self-identification, and the question should allow respondents to select multiple ethnicities. While there are several methods by which multiple ethnic identifications can be collected (e.g., check-all-that-apply on a predefined list, selection of a specific ethnic combination, and an open-ended text box), and no one-size-fits-all approach, there is argument that check-all-that-apply is likely the optimal method for the CANZUS states and United Kingdom countries (Aspinall, 2018b; Balestra & Fleischer, 2018; Morning, 2008; United Nations, 2008).

Conversely, substantially less attention has been directed towards how researchers should classify multiple ethnic identifications for quantitative analysis, and why these methodological decisions matter. Ethnic classification refers to preparing raw ethnicity data collected from surveys (input) into a format suitable for statistical analysis and reporting (output). Multiple ethnicity data can present challenges to applied researchers as many common statistical techniques typically require mutually exclusive groups with adequate sample sizes (e.g., chi-square test of independence, independent samples *t*-test, ANOVA, regression analyses, and structural equation modelling [SEM]; Cohen et al., 2003; Field et al., 2012). As noted in Section 1.2, while there is relatively little guidance in the literature on how to output these data, there is a small body of work that addresses this issue, particularly from the United States (e.g., Liebler & Halpern-Manners, 2008; Mays et al., 2003; OMB, 2000) and Aotearoa New Zealand (e.g., Callister et al., 2007; Cormack & Robson, 2010; Statistics New Zealand, 2004). The

majority of this literature is published between 2000 and 2010, reflecting heightened research and policy interest in multiple ethnic identifications at the time (Morning, 2008). Less has been written on this topic since then, although it is arguably even more pertinent now, given the steadily increasing rate of multiple ethnic identifications around the world (Aspinall, 2018b; N. Jones et al., 2021; Statistics New Zealand, 2014b). The following section provides a high-level overview of the ethnic classification literature (see Section 4.1.2 for a more detailed discussion).

### 2.3.2 Issues in Ethnic Classification

Ethnic classification methods described in the literature are largely based on official government documents on statistical standards for ethnicity. In the United States, the key relevant document is the OMB's (2000) *Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity*, which describes several classification methods to bridge the transition from federal collection of single racial responses (original 1977 standards) to multiple racial responses (revised 1997 standards). The methods are split into three broad approaches:

- *Whole assignment*, where multiple responses are assigned to a single category based on a deterministic assignment scheme (usually according to relative population size, e.g., "Smallest Group", "Largest Group", "Largest Group Other Than White"), or a probabilistic distribution;
- *Fractional assignment*, where multiple responses are assigned to multiple categories using fractional weightings that sum to 1 for each respondent; and
- The *all-inclusive method*, where respondents are assigned to each of the categories they select, such that the sum of categories can exceed the number of respondents.

Some U.S. researchers also used data from the National Health Interview Survey (NHIS), which asked multiracial participants to select all their racial groups, as well as a

follow-up question about the race they identified with most ("main race"), to develop the *NHIS regression method.* This method uses covariates such as age, sex, region, urbanicity, and area racial composition to predict the "main" race for each multiracial combination (Ingram et al., 2003; Liebler & Halpern-Manners, 2008; Schenker & Parker, 2003). An alternative method described by the OMB (2000), which results in a break in time series reporting when the option of multiple racial identifications was introduced, is the classification of these responses into *specific racial combinations* (e.g., Black/White, Asian/White, Black/Asian).

In Aotearoa New Zealand, the key relevant document is Statistics New Zealand's (2004) *Report of the Review of the Measurement of Ethnicity*, which outlines four ethnic classification methods to inform recommendations for Statistics New Zealand's (2005) *Statistical Standard for Ethnicity*. The methods considered were:

- *Total response output*, where respondents are assigned to each of the categories they select (i.e., ethnic group counts can exceed the number of respondents);
- *Sole/combination output*, where respondents are assigned to a mutually exclusive category that indicates their sole ethnic group or specific ethnic combination;
- *Prioritised ethnicity*,[5] where multiple responses are assigned to a single category based on a predetermined hierarchy (Māori > Pacific Peoples > Asian > Other > European); and
- *Self-prioritised ethnicity*, where multi-ethnic respondents are asked to select their "main" ethnic group.

It is evident that, despite differences in terminology, there is considerable overlap in racial/ethnic classification methods between the two countries. Section 4.1.2, which discusses ethnic classification methods in more detail, presents an organising framework for these

[5] In this thesis, prioritised ethnicity is referred to as "administratively-prioritised ethnicity" to avoid potential confusion with self-prioritised ethnicity.

methods. Advantages and disadvantages of each method are reviewed in the section, showing that there is no clear "best" method. This raises the key points of interest in this thesis: which ethnic classification method researchers should use, whether their decision will make a substantive difference towards results, and if so, what effects the decision will have.

The small body of existing research in this area from the United States and Aotearoa New Zealand indicates that, as expected, ethnic classification method affects the absolute and relative sizes of ethnic groups (Callister et al., 2007; Grieco, 2002; Parker & Makuc, 2002). For example, Grieco's (2002) analysis of 2000 U.S. Census data, where 2.4% reported more than one race, showed that estimates for the Native Hawaiian/Other Pacific Islander grouping (numerically smallest group) varied by over a factor of 2 depending on the classification method (estimates ranged from 399,000 using the Largest Group deterministic scheme, which prioritises largest groups first; to 874,000 using the Smallest Group deterministic scheme, which prioritises smallest groups first). This has important implications because ethnic group size impacts ethnic group recognition, statistical representation, and resourcing decisions (Balestra & Fleischer, 2018; Callister et al., 2007; Mays et al., 2003). In addition, when calculating prevalence statistics, inconsistent methods used to output the numerator and denominator can lead to the numerator-denominator bias, where results (e.g., mortality rates by ethnicity) are inaccurate due to the different methods used (Balestra & Fleischer, 2018; Blakely et al., 2005; Callister et al., 2007; Mays et al., 2003).

Furthermore, existing research in these countries show that even when the same method is used for the numerator and denominator, prevalence statistics can differ depending on the ethnic classification method used. For example, Mays et al.'s (2003) analysis of data from the 2001 California Health Interview Survey (5% multiracial) showed that the proportion of Native Hawaiian/Other Pacific Islander adults who reported no usual source of healthcare ranged from 17% to 22% depending on the classification method. There is also evidence that, particularly

for numerically smaller groups, ethnic classification method affects prevalence statistics in asthma, diabetes, and heart disease diagnoses (Mays et al., 2003; [New Zealand's] Ministry of Health, 2008); health insurance status (Mays et al., 2003; Parker & Makuc, 2002); low birthweight (Heck et al., 2003); hospitalisation with infectious disease (Hobbs et al., 2019); smoking rates (Boven et al., 2020); mortality rates (Callister & Blakely, 2004); and so on. These differences in prevalence outcomes, which are due solely to ethnic classification differences, can impact conclusions about existing ethnic inequities, as well as decisions around targeted interventions and policies.

However, there are several gaps in the literature. First, self-prioritisation is seldom included in studies that investigate the effects of different ethnic classification methods on quantitative analysis (these studies usually examine total response, sole/combination grouping, and administrative-prioritisation). Thus, the empirical implications of using self-prioritisation is relatively unknown. Second, the majority of existing research was conducted with adult populations with less than 10% multi-ethnic prevalence (Didham & Callister, 2012; Liebler & Halpern-Manners, 2008; Mays et al., 2003; Ministry of Health, 2008; Te Rōpū Rangahau Hauora a Eru Pōmare, 2000). As the global multi-ethnic population continues to increase, particularly among children and adolescents, it is important to investigate how ethnic classification method will affect quantitative research involving higher multi-ethnic identification rates. Third, as implied by systematic reviews which highlight the lack of transparency in ethnicity reporting as a major flaw in existing research (Bokor-Billmann et al., 2020; Craven et al., 2018; Maduka et al., 2021; Moore, 2020; Moubarac, 2013), a large number of researchers are either unaware of, or do not explicitly consider, the potential impact of ethnic classification decisions on research. Therefore, there is a clear need for further research which empirically investigates the effects that ethnic classification method has on statistical analysis and resulting conclusions in increasingly multi-ethnic contexts, in order to inform researchers

about the methodological and substantive implications of their decisions on the important variable of ethnicity.

## 2.4 Aotearoa New Zealand Context

Aotearoa New Zealand is a valuable context for research on ethnic classification because of its ethnically diverse population and relatively high rates of multi-ethnic identification (Balestra & Fleischer, 2018; Statistics New Zealand, 2014b, 2020a). It was one of the early countries to adopt a self-identified definition of ethnicity and collect multiple ethnicity data in official statistics (Balestra & Fleischer, 2018; Callister et al., 2007; Cormack & Robson, 2010), and was recognised as demonstrating a "good practice example" (p. 37) of ethnicity statistical standards in Balestra and Fleischer's (2018) review of diversity statistics in countries in the Organisation for Economic Co-operation and Development (OECD). In addition, Article 3 of Aotearoa New Zealand's founding document, *Te Tiriti o Waitangi* (te reo Māori [Māori language] version of *The Treaty of Waitangi*), guarantees Māori the same rights and privileges of British subjects (Orange, 2011). In the modern context, this means the Crown is obliged to reducing existing ethnic disparities and achieving equity in outcomes (Kingi, 2007; Waitangi Tribunal, 2019; Wyeth et al., 2010). This agenda is also of important interest to researchers in Aotearoa New Zealand, making ethnic classification a pertinent issue in research and policy (Atatoa Carr et al., 2017; Callister et al., 2007).

Ethnic classification in each country is inextricably linked to its sociohistorical and political context (Balestra & Fleischer, 2018; Morning, 2008; Rocha & Aspinall, 2020; United Nations, 2008), so Aotearoa New Zealand's sociohistorical background will be outlined next to contextualise the thesis. There will be a focus on ethnicity-related official census practices and statistical standards, as these reflect the socio-political climate at the time, and also inform survey design (Balestra & Fleischer, 2018; Kertzer & Arel, 2002; Rocha & Aspinall, 2020).

Finally, current ethnic classification practices in research in Aotearoa New Zealand will be reviewed.

### 2.4.1 Sociohistorical Background

Aotearoa New Zealand is a colonial settler state in the South Pacific first inhabited by Māori, the Indigenous Peoples of the land (tangata whenua), around 1300 CE. European settlers/colonisers began arriving in the late 1700s. The majority of early arrivals were male whalers, sealers, and traders; and there was intermarriage between Māori and non-Māori from early European contact (Callister et al., 2007; Khawaja et al., 2000). In 1840, *Te Tiriti o Waitangi*, the country's founding document comprising three articles relating to Crown governance (kāwanatanga), Māori chieftainship (tino rangatiratanga) over lands and treasured possessions, and the same protections and rights (ōritetanga) for Māori as that accorded to British subjects, was signed between the British Crown and Māori chiefs (Orange, 2011). Unlike other CANZUS states, intermarriage was never formally restricted in Aotearoa New Zealand, but was instead seen as a way for European colonisers to biologically and culturally assimilate Māori, as well as acquire their land (Cormack & Robson, 2010; Rocha, 2012). Nonetheless, there was negative public discourse and attitudes toward "miscegenation" as this disrupted the status quo (Rocha, 2012). From the mid-1900s, Aotearoa New Zealand experienced several significant waves of migration—first from neighbouring Pacific nations (e.g., Sāmoa, Tonga, and Fiji) due to demand for labour; then from East Asian, Southeast Asian, and South Asian countries (e.g., China, India, Philippines) following a relaxing of immigration policies to attract skilled and business migrants from Asia (note a small number of Chinese workers had also settled earlier during the 1860s Gold Rush); and subsequently from the Middle East and Africa (e.g., Iran, Iraq, Somalia), many as refugees (Callister et al., 2007; Khawaja et al., 2000).

Enumeration in early colonial censuses was based on notions of race, "blood quantum", and "proportion-of-descent" (e.g., "half-caste" and "quarter-caste"). Up until 1974, only those who reported at least half Māori descent were classified as Māori, reflecting the state's agenda of assimilation (Cormack & Robson, 2010; Rocha, 2012). This longstanding colonial practice was contradictory to Māori conceptualisation of group belonging, which is based on whakapapa (ancestry) that cannot be divided into parts (Jackson, 2003; Khawaja et al., 2000; A. D. Williams et al., 2018). However, following Māori rights movements and increasing opposition against the concept of race, significant changes in the five-yearly census were later introduced. In 1976, the census question changed from enumerating "race" to "ethnic origin", and an additional question on Māori ancestry was included; in 1986, proportion-of-descent was replaced with a check-all-that-apply response format; and in 1991, "ethnic origin" was modified to "ethnic group" (Cormack & Robson, 2010; Khawaja et al., 2000). From 1991, the standard ethnicity census question in Aotearoa New Zealand is: "Which ethnic group do you belong to? Tick the box or boxes which apply to you" (Cormack & Robson, 2010; Khawaja et al., 2000). This remains the current official wording and is also frequently used in sample surveys and questionnaires (Atatoa Carr et al., 2017; Clark et al., 2013).

The most recent 2018 Census enumerated a usually resident population of 4.7 million people in Aotearoa New Zealand. Seventy percent identified with a European ethnic group; 17% identified with the Māori ethnic group; 15% identified with an Asian ethnic group;[6] 8% identified with a Pacific ethnic group; 1.5% identified with a Middle Eastern, Latin American, or African (MELAA) ethnic group; and 1.2% identified with an ethnicity classified to the

[6] Note the official Statistics New Zealand (2005) definition of "Asian" excludes most ethnicities originating from West Asia (e.g., Iranian, Iraqi, and Syrian). These are instead included in the MELAA category.

residual "Other" category (Statistics New Zealand, 2020a).[7] The sum of these percentages exceed 100% because respondents could select ethnicities in more than one of these broad ethnic groupings.

Overall, 11% of the population reported more than one broad ethnic grouping, but there was notable variation by ethnicity and age group (Statistics New Zealand, 2020a). Specifically, multiple ethnic identifications were substantially more common in those who identified with a Māori (55%) or Pacific ethnic grouping (41%), compared to those who identified with a European (16%), Asian (10%), MELAA (22%), or Other (20%) ethnic grouping (Statistics New Zealand, 2020a). Note the broad ethnic groupings reported here are utilised for statistical output only. There is considerable intra-group heterogeneity, and multi-ethnic identification can also occur within broad ethnic groups (e.g., a person can identify as both Sāmoan and Tongan, but will not be officially counted as multi-ethnic when broad ethnic groupings are used; Atatoa Carr et al., 2017). In addition, multi-ethnic identification was more common among younger age groups (23% in those aged under 15 years, 15% in those aged between 15 and 29, 8% in those aged between 30 and 64, and 3% in those aged 64 or over; Statistics New Zealand, 2020a). Note multi-ethnic rates tend to be lower in population censuses than sample surveys—for example, nationally representative surveys of children and adolescents used in this thesis show that more than 30% report more than one broad ethnic group (Atatoa Carr et al., 2017; Clark et al., 2013; Fleming, Peiris-John, et al., 2020). This can be due to a variety of reasons, including simplification of responses in official forms, where and how the question was asked, as well as beliefs about how the data may be used (Callister et al., 2007).

[7] Previous research indicates that the majority in the "Other" category identify as "New Zealander" and are of European descent (Broman, 2018; Kukutai & Didham, 2009).

### 2.4.2 Current Statistical Standard for Ethnicity

Statistics New Zealand's (2005) *Statistical Standard for Ethnicity* was developed to encourage a standardised definition of ethnicity in Aotearoa New Zealand, as well consistent data collection, coding, and output.[8] Ethnicity is defined in the standard as follows:

> Ethnicity is the ethnic group or groups that people identify with or feel they belong to. Ethnicity is a measure of cultural affiliation, as opposed to race, ancestry, nationality or citizenship. Ethnicity is self perceived (sic) and people can belong to more than one ethnic group. (Statistics New Zealand, 2005, p. 1)

This definition has two main implications for standard ethnicity data collection: (1) response options should allow respondents to select their ethnic group(s) without being forced into a more general category, and (2) the form should allow respondents to select more than one ethnic group. To facilitate standardised ethnicity response options, a four-level hierarchical ethnic classification system of increasing specificity is provided in the standard:

- *Level 1* has 6 categories (European, Māori, Pacific Peoples, Asian, MELAA, and Other), and is used for statistical output only;
- *Level 2* has 21 categories (e.g., the European category is divided into New Zealand European and Other European; the Pacific Peoples category is divided into Sāmoan, Cook Islands Māori, Tongan, Niuean, etc.);
- *Level 3* has 36 categories (e.g., the Other European category is divided into British and Irish, Dutch, Greek, Polish, etc.); and
- *Level 4* has 180 categories (e.g., the British and Irish category is divided into British, English, Irish, Scottish, Welsh, etc.).

[8] Statistics New Zealand (2020b) conducted a public consultation of the ethnicity standard in late 2019, and has signalled future work with users, stakeholders, and subject experts.

The ethnicity standard states that, at a minimum, ethnicity should be collected at Level 2 of the classification system. However, for statistical output, responses can be aggregated into broader categories as needed (e.g., at Level 1).

In the ethnicity standard, Statistics New Zealand (2005) also lists two standard ethnicity output methods:

- *Total response output*, where respondents are counted in each of their ethnic groups (e.g., European, Māori, Pacific Peoples, Asian, MELAA, and/or Other); and
- *Sole/combination output*, where respondents are counted once in their sole ethnic group (e.g., European only, Māori only, Pacific Peoples only, etc.) or combination of ethnic groups (e.g., Māori/European, Māori/Pacific Peoples, Pacific Peoples/European, etc.[9]).

From the 63 permutations possible at Level 1, three standard sets of increasingly detailed sole/combination categories are outlined, and includes 8 groups (recommended for small datasets), 15 groups (recommended for large datasets), and 45 groups (recommended for very large datasets), respectively. Endorsement of total response and sole/combination output was based on recommendations in the *Report of the Review of the Measurement of Ethnicity* (Statistics New Zealand, 2004). These methods were recommended because they retain multiple ethnic identifications by respondents.

The *Report of the Review of the Measurement of Ethnicity* (Statistics New Zealand, 2004) also recommended the discontinuation of the previously standard output method of administrative-prioritisation (where multiple responses are classified into single categories according to Department of Statistics' [1993] hierarchy:[10] Māori > Pacific Peoples > Asian > Other > European). Reasons included its biasing effect on statistics, especially as multi-ethnic

[9] In this thesis, ethnic combinations are ordered according to Department of Statistics' (1993) prioritisation hierarchy (i.e., Māori/European, rather than European/Māori).

[10] The Department of Statistics was renamed to Statistics New Zealand in 1994.

identifications increase; and inconsistencies with the official definition of ethnicity (i.e., ethnicity is self-perceived and can include multiple affiliations). Similarly, self-prioritisation (where respondents are asked to select their "main" ethnic group) was considered but not recommended in the report because of inconsistencies with the definition of ethnicity.

### 2.4.3 Ethnic Classification in Research

Although Statistics New Zealand's (2005) endorsement of two standard output methods (total response and sole/combination output) may make ethnic classification appear to be a fairly straightforward process, in reality it is a complex issue in applied research. A number of scholars in Aotearoa New Zealand have written about the relative strengths and limitations of each of the ethnic classification methods (Atatoa Carr et al., 2017; Callister et al., 2007; Cormack & Robson, 2010; Didham, 2005). For example, a main strength of total response and sole/combination output is the retention of multi-ethnic identifications, but the former counts multi-ethnic individuals multiple times such that they can be over-represented in analyses, and the latter can result in a large number of categories—some with very small subgroup sizes and thus unequal explanatory power. On the contrary, although the prioritisation methods' main weakness is their inconsistency with the official definition of ethnicity, administrative-prioritisation can be advantageous in ensuring that numerically smaller and marginalised groups are not subsumed in the data, and self-prioritisation can indicate the ethnic group that multi-ethnic individuals identify most strongly with, if they have one.

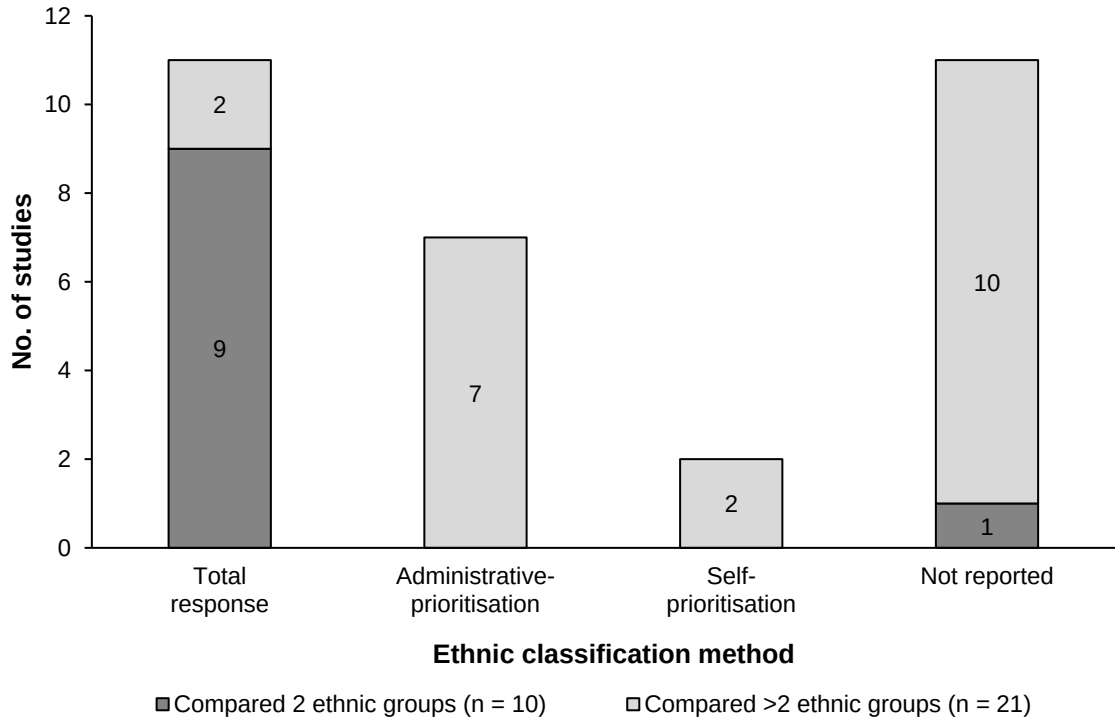
In the academic literature, scholars in Aotearoa New Zealand and internationally broadly agree that, when working with ethnicity data, researchers should choose the ethnic classification method that is most appropriate for their specific research question and context, regardless of whether the method is officially endorsed (Atatoa Carr et al., 2017; Cormack & Robson, 2010; Moubarac, 2013; Woo et al., 2011). Meanwhile, there are also growing calls for Indigenous data sovereignty in Aotearoa New Zealand and other colonised states, whereby

Indigenous Peoples determine how data involving Indigenous Peoples are collected, analysed, interpreted, and disseminated (Cormack et al., 2019; Kukutai & Taylor, 2016). Naturally, this includes decisions on how ethnicity data are classified for statistical analysis, and is particularly important because data are not neutral, but are subjective, political, and have typically been used (both intentionally and unintentionally) in ways that oppress Indigenous Peoples (Smith, 2021; Walter & Andersen, 2013). As mentioned in Section 2.3.2, there is some empirical research in Aotearoa New Zealand, typically using census data or large-scale surveys with adults, which show the impact that ethnic classification method can have on ethnic group counts (Callister et al., 2007; Didham & Callister, 2012), as well as on interpretations of ethnic disparities in health and social outcomes (Boven et al., 2020; Callister & Blakely, 2004; Curtis et al., 2005; Ministry of Health, 2008; Te Rōpū Rangahau Hauora a Eru Pōmare, 2000). Together, this underscores the complex but important decision of ethnic classification method in quantitative research, the need for researchers to honour Indigenous data sovereignty, as well as the importance of being transparent about the ethnic classification method used, as these can all affect research conclusions.

However, similar to international patterns (Moubarac, 2013), there appears to be a gap between the aforementioned literature on ethnic classification and common practice in quantitative research in Aotearoa New Zealand. To scope common practices in the country, I conducted an exploratory systematic PsycINFO database search of articles published between 2018 and 2020 which contained the words “health”, “ethnic*”, and “New Zealand” in the title or abstract (the asterisk finds terms starting with the letters “ethnic”, i.e., both “ethnic group” and “ethnicity” will be included). Figure 2.1 shows the ethnic classification method used in quantitative studies that compared at least two ethnic groups. Of the 31 studies identified, only 65% (20/31) reported the ethnic classification method used, or provided enough information that it could be deduced. This transparency rate is somewhat higher than that reported in

**Figure 2.1**

*Ethnic Classification Methods Used in Health Studies in Aotearoa New Zealand Published Between 2018 and 2020 (N = 31)*




Moubarac's (2013) systematic review (58%), which mainly consisted of U.S. articles, but is nonetheless still concerning.

Interestingly, as shown in Figure 2.1, the exploratory database search suggests that the transparency rate, as well as the ethnic classification method used in each study, differ depending on the number of ethnic groups compared. Specifically, of the studies that compared two ethnic groups, nearly all of these specified the ethnic classification method used, or provided enough information for this to be deduced (9/10). Total response (e.g., Māori vs. non-Māori) was used to compare two "ethnic groups" [11] in all the studies with available information

[11] Note that, strictly speaking, "non-*x*" (e.g., non-Māori) is not an ethnic group, but a category containing residual responses.

(9/9). On the contrary, of the studies that compared more than two ethnic groups, just over half reported the ethnic classification method used, or provided enough information for this to be inferred (11/21). Where reported, administrative-prioritisation was the most common method used to compare more than two ethnic groups (7/11), followed by total response (2/11) and self-prioritisation (2/11). Based on my observations and experience in completing this thesis, of the studies in Aotearoa New Zealand that did not report the ethnic classification method used, it is likely that administrative-prioritisation was used with administrative datasets, and self-prioritisation with primary datasets.

This exploratory systematic database search shows that despite the small body of existing literature highlighting the importance of ethnic classification decisions in quantitative research, many researchers—especially when engaging in analyses with more than two ethnic groups—are not sufficiently transparent about ethnic classification. As Moubarac (2013) indicated, it is likely that this lack of transparency is due to many mainstream researchers either being unaware of the complexities regarding ethnic classification, or not fully realising the implications these decisions can have on research. This is increasingly of concern as multi-ethnic identification becomes more common (Atatoa Carr et al., 2017; Statistics New Zealand, 2014b, 2020a).

## 2.5 The Current Study

Ethnicity is a variable of research and policy importance due to ongoing ethnic inequities, but the operationalisation of ethnicity is complex due to increasing rates of multi-ethnic identification. Existing research, primarily conducted utilising adult samples with relatively low multi-ethnic prevalence (e.g., <10%), indicates that researchers' choice of ethnic classification method can influence research findings. However, there are three key gaps in this field: (1) there is scant research that investigates the effects of ethnic classification method on

samples with higher multi-ethnic prevalence (e.g., >10%), (2) the effects of self-prioritisation on data analysis are underexplored, and (3) there is a concerning lack of transparency in ethnicity reporting in published articles. Therefore, the current thesis aims to empirically explore the impact of ethnic classification method on quantitative analysis with increasingly multi-ethnic datasets (adults, adolescents, and children) in order to inform researchers and policymakers of the potential implications of their methodological decisions. Data from the Growing Up in New Zealand study will be utilised for the child and adult samples, and Youth'12 data will be utilised for the adolescent sample. The overarching research question of this thesis is: *how does ethnic classification method affect applied quantitative analysis in multi-ethnic contexts?* This research is important for equity and social justice, because ethnicity is not a neutral and objective variable, but a subjective and political one that has frequently been used to marginalise Indigenous and ethnic minority groups.

Because the current research is intended to be relevant to as many applied researchers as possible, the effects of ethnic classification method will be explored on cross-sectional, self-identified ethnic groups when aggregated into broad ethnic groupings (e.g., European, Māori, Pacific, Asian, MELAA, Other)—the most common way of collecting and processing ethnicity data for quantitative analysis in Aotearoa New Zealand (Statistics New Zealand, 2005). While there is heterogeneity within each ethnic grouping, it nonetheless is a useful starting point given the exploratory nature of this research. The four most accessible ethnic classification methods will be investigated: total response grouping, sole/combination grouping, administrative-prioritisation, and self-prioritisation. The first two methods are both recommended by Statistics New Zealand (2005), whereas the third and fourth methods are more commonly used in health and social science research.

The implications of researchers' ethnic classification decisions are investigated in three empirical studies with increasing specificity. These studies are described below, and

summarised in Table 2.1. Study 1 investigates response patterns to the self-prioritisation question in children (via mother proxy), adolescents, and adults; as well as how the three classification methods that categorise multiple ethnicity data into a small number of broad ethnic groups (i.e., total response, administrative-prioritisation, and self-prioritisation) affect outputted ethnic group sizes in each age cohort. This is a descriptive study that scopes the extent to which ethnic classification method affects ethnicity as a variable in increasingly multi-ethnic cohorts, before other demographic variables and outcome measures are incorporated.

Study 2 focuses on the most commonly used ethnic classification methods of administrative-prioritisation and self-prioritisation, and investigates how discrepancies between these two methods in each age cohort are associated with individual characteristics (e.g., ethnic combination, age, sex, birthplace, and education level) and contextual characteristics (e.g., urbanicity, neighbourhood socioeconomic deprivation, and neighbourhood ethnic composition). This correlational study examines how ethnic classification method influences the demographic composition of independent variables that are frequently included with ethnicity in statistical analysis, and thus, the extent that ethnic

**Table 2.1**

*Overview of this Thesis's Studies*

| | Study 1 | Study 2 | Study 3 |
|---|---|---|---|
| Research focus | • Outputted ethnic group size | • Demographic composition | • Mental health outcomes |
| Quantitative design | • Descriptive | • Correlational | • Descriptive<br>• Correlational |
| Age cohorts | • Children<br>• Adolescents<br>• Adults | • Children<br>• Adolescents<br>• Adults | • Adolescents |
| Datasets | • Growing Up in NZ<br>• Youth'12 | • Growing Up in NZ<br>• Youth'12 | • Youth'12 |

classification method biases analytic samples in increasingly multi-ethnic age cohorts.

Study 3 explores how ethnic classification method affects substantive mental health outcomes, both within ethnic groups (e.g., absolute prevalence rates for each ethnic group) and between ethnic groups (e.g., relative difference between ethnic groups). In order to ascertain whether ethnic classification effects are similar across outcomes, three mental health outcomes will be investigated: overall psychosocial difficulties, deliberate self-harm, and suicide attempt. Due to its focus on these mental health outcomes, the study will only be conducted with the adolescent sample (see more detailed explanation in Chapter 6). All four common ethnic classification methods will be examined: total response, sole/combination grouping, administrative-prioritisation, and self-prioritisation. This study can be characterised as being both descriptive and correlational, because although its aim is to describe the differences that ethnic classification method has on outcomes, it employs the correlational technique of regression so that common demographic characteristics (e.g., age, sex, urbanicity, and socioeconomic deprivation) can be incorporated as independent variables alongside ethnicity. Thus, Study 3 brings together the elements examined in the previous two studies—namely ethnic group size and demographic composition—and adds an outcome dimension to it.

Together, the three studies are intended to address the aforementioned gaps in the literature, and comprehensively inform researchers and policymakers of the implications of ethnic classification decisions.

CHAPTER 3

# Methodology: Positionality Matters

Chapter 2 has situated the current thesis in the literature, argued for why ethnicity matters, and established the research gaps. It also outlined how these gaps will be addressed through three studies that will inform researchers of the implications of ethnic classification method on outputted ethnic group size, demographic composition, and substantive outcomes, respectively. The current chapter describes the overarching methodology of the thesis, starting with my positionality as a researcher and the research paradigm this thesis is situated within. Then, the data sources utilised in this thesis are introduced, and the main ethical considerations of the research are discussed. Specific methods used in each of the three studies will be described in their respective chapters.

Mainstream quantitative researchers tend to not state their positionality and research paradigm, because quantitative research is often considered to be conducted from a "neutral" and "objective" perspective (Lincoln et al., 2018). However, especially given the subjective and political nature of ethnicity data, it is important for researchers to make their positionality explicit as this influences the entire research process—from what research questions are prioritised, to what data are collected, how ethnicity and other variables are processed and analysed, and particularly, how the results are interpreted (Bonilla-Silva & Zuberi, 2008; Gillborn et al., 2018; Walter & Andersen, 2013). As Zuberi writes, "The views and social position of researchers have a lot to do with how they interpret racial statistics. . . . Data do not tell us a story. We use data to craft a story that comports with our understanding of the world" (Bonilla-Silva & Zuberi, 2008, p. 7).

## 3.1 Researcher Positionality

As stated in the opening paragraph of Chapter 1, I embarked on the research topic of ethnic classification because I wanted to determine what ethnic classification method was best to use in my research, and how this can influence my results. I entered into this process with an open mind and no preconceptions about which method might be better. However, I acknowledge that my positionality, particularly my insider/outsider status on the key issues in this thesis, influences my search for answers (Finlay, 2002). Rather than approaching insider/outsider status from a dichotomous perspective, I consider myself as occupying the "the space between" (Dwyer & Buckle, 2009) in the four main themes of this thesis: ethnic identification, multi-ethnicity, colonisation, and racism.

I was born in British Hong Kong, was four years old when Hong Kong's sovereignty was returned to the People's Republic of China (1997), and seven years old when my family and I immigrated to Aotearoa New Zealand (2000). In Aotearoa New Zealand, my parents strove to preserve and continue to instil Chinese language, culture, and values in me; while I endeavoured to fit in (assimilate) to dominant New Zealand culture in every way I could. Due to this background, I prefer to self-identify as a 1.5-generation Asian-New Zealander, as this best captures my feelings of being in-between two cultures. In terms of ethnic identification, when asked the ethnicity question, I typically tick the "Chinese" box because I feel it is the most acceptable response. However, this is usually accompanied with an uncomfortable feeling. First, "Chinese" (中國人) is incongruent with my private ethnic identification—like many Hong Kong-born Chinese, I tend to prefer self-identifying with the panethnic group of "Asian"

(亞洲人) or the local identity of "Hongkonger" (香港人).[12] In addition, none of these identifications acknowledge the hybridity I feel as a 1.5-generation Asian-New Zealander. An open-ended "other—please state" response option is usually provided, but influenced by the Chinese value of respecting authority (Mok & Defranco, 2000), I typically do not have the courage to be "defiant" enough to select this option. The other reason I feel uncomfortable with being asked the ethnicity question is because it makes me feel "outed" as an outsider and perpetual foreigner in Aotearoa New Zealand. Therefore, although ethnicity data have usually not been used to frame Asian Peoples in a deficit manner (the way it has unfortunately been used in relation to Māori and Pacific Peoples), I understand and am sensitive about the tensions and reservations the ethnicity question can invoke.

In terms of the theme of multi-ethnicity, I do not have lived experience as a multi-ethnic individual. However, I have Vietnamese heritage through my great-grandmother, and have grown up being told that "that makes you one-eighth Vietnamese". This contributed to my initial thinking that self-prioritisation is a promising ethnic classification method because it can elicit the ethnicity a person identifies with most (this was before I fully understood the distinction between ancestry and ethnicity). It was only through subsequent readings (e.g., Aspinall & Song, 2013; Jackson, 2003; Rocha & Webber, 2018; Sanchez, 2010; Townsend et al., 2009; among many others), seminars, and conversations with people with lived multi-ethnic experience, that I began to understand the problems with "forcing" multi-ethnic individuals to prioritise one of their ethnic groups. My lived experience as a 1.5-generation Asian-New Zealander feeling caught between two cultures helped me be open and empathetic to these

[12] Results from the Hong Kong University Public Opinion Programme's (2019) biannual survey showed that around three-quarters of local respondents preferred to identify as "Hongkongers" (i.e., "Hongkongers" or "Hongkongers in China"; 76%) rather than "Chinese" (i.e., "Chinese or "Chinese in Hong Kong"; 23%).

perspectives, leading to a re-evaluation of my initial bias towards self-prioritisation, and heightened reflexivity of how my background influences my views on each ethnic classification method.

My position regarding colonisation is also something that has changed throughout the course of my PhD. Hong Kong, like Aotearoa New Zealand, was a British colony, and I was exposed to notions of Western superiority from an early age. However, there is an important distinction: on balance, Hongkongers tend to perceive British colonisation as a positive process that brought prosperity, democracy, freedom of speech, and civilisation to the city (Luk, 2017). In addition, British Hong Kong provided a safe place for my grandparents to flee to as refugees during the Chinese Civil War. These influences led to my initial crude perspective that colonisation is beneficial, and admittedly, some difficulty truly understanding why Māori typically view colonisation of Aotearoa New Zealand negatively. Ironically, it was developments in post-colonial Hong Kong (e.g., reductions in democracy and freedom of speech, civic education of national [Chinese] identity over local [Hong Kong] identity, and promotion of the Mandarin Chinese dialect over the local Cantonese dialect)—viewed by many Hongkongers as "recolonisation" or "motherland colonisation" (Luk, 2017)—that helped me see colonisation with more nuance. I now recognise that my initial perspective towards colonisation was specific to the British Hong Kong context, understand the importance of honouring *Te Tiriti o Waitangi*, and position myself as an Asian tangata Tiriti and ally to Indigenous Peoples.

Finally, as a member of an ethnic minority group living in Aotearoa New Zealand, I have experienced racism, mostly through micro-aggressions that emphasise I am different and do not belong in the country. I have come to realise that I also struggle with internalised racism, which in the context of this thesis, primarily manifests as frequent questioning about my place in researching and contributing to the topic of ethnic classification in Aotearoa New Zealand.

In contrast, my awareness of experiences around structural racism is more limited. However, as a result of my readings and conversations with others about various levels of racism, and aided by my position as an ethnic minority, I now comprehend the impact of racism in more depth, including how experiences of racism can differ between (and within) ethnic groups. I now also recognise the importance of struggling for anti-racism, equity, and social justice—both at the individual and societal levels.

In summary, I identify my positionality as being in the space between an insider and outsider in my research. This positionality enables me to engage in this research with a unique mix of nuanced understanding and open-mindedness. In addition, I am highly aware of how I am likely to be perceived as an outsider in the field of ethnic classification in Aotearoa New Zealand. Thus, I am particularly reflexive about how my position as an Asian immigrant influences the way I approach my thesis and interpret my findings, and how I can be an ally to tangata whenua (people of the land, i.e., Māori) as an Asian tangata Tiriti.

## 3.2 Research Paradigm

The research paradigm I approach this thesis with is also important to make explicit, because this influences my perspectives about ontology (nature of reality), epistemology (relationship between the researcher and what is being researched), methodology (how knowledge is obtained), and axiology (values guiding the research; Creswell, 2013). I initially developed and approached my research questions under a post-positivist paradigm, which assumes that a single objective reality exists, but that this reality can only be known imperfectly (Lincoln et al., 2018).[13] This is reflected in the main question that initially motivated this

[13] Post-positivism was developed as a critique of positivism. Positivists operate under the assumption that there is an objective, apprehensible reality (Lincoln et al., 2018).

thesis—that is, "which ethnic classification method is closest to approximating reality?" However, I shifted towards a critical paradigm over the course of this PhD as a result of my many readings and conversations with others. Critical paradigms (e.g., critical race theory, feminist theory, and queer theory) emphasise the role that socio-political dynamics—particularly pertaining to interactions between the privileged and the oppressed—shape reality, and are concerned with producing research that results in positive change for those oppressed by power (Kincheloe et al., 2018; Lincoln et al., 2018). While this thesis is written from a critical paradigm, it also contains elements of constructivism (e.g., what is "real" depends on the person constructing the reality) and pragmaticism (e.g., inquiry methods should be selected based on the research purpose; Mertens, 2019).

One of the main differences between post-positivist and critical paradigms is that the goal of post-positivist researchers is to objectively conduct research that most closely approximates reality (Lincoln et al., 2018). This is typically done using quantitative methods. Under post-positivism, involvement in action resulting from research is considered a form of subjectivity which undermines researcher objectivity (Lincoln et al., 2018). In comparison, the goal of critical researchers is to conduct research that results in action towards social change—particularly equity, social justice, and emancipation. This is typically done using qualitative methods which give voice to the oppressed (Lincoln et al., 2018). Although qualitative data are often viewed as being better suited for exposing and understanding the nuances of oppressive social processes, a number of scholars have argued that quantitative data are not inherently antithetical to critical research agenda (Bonilla-Silva & Zuberi, 2008; Cokley & Awad, 2013; Garcia et al., 2018; Gillborn et al., 2018; Sablan, 2019; Stage, 2007; Stage & Wells, 2014). For example, Stage (2007) argues that critical paradigms are distinguished from post-positivist paradigms by the motivation of the research, rather than the specific methods used (e.g.,

qualitative vs. qualitative); and Gillborn et al. (2018) argues that, when used appropriately and reflexively, quantitative methods can contribute to social justice.

Therefore, aligning with its quantitative nature, the current thesis is situated within a critical quantitative paradigm. Originally developed in the context of research in higher education, critical quantitative inquiry involves questioning and challenging conventional concepts, measures, and processes, with the goals of equity and social justice (Rios-Aguilar, 2014; Stage, 2007; Stage & Wells, 2014). The three main tasks of critical quantitative inquiry as outlined by Stage and colleagues are to: (1) use quantitative data to identify inequitable processes and outcomes; (2) question methodological practices in quantitative research, with the aim of better representing those who are marginalised; and (3) conduct culturally relevant research by considering people in their wider institutional contexts (Stage, 2007; Stage & Wells, 2014). The specific focus of this thesis is to critically examine ethnic classification practices and how these affect the processes and outcomes of quantitative research.

## 3.3 Data Sources

This thesis utilises secondary datasets from two large-scale research studies in Aotearoa New Zealand to investigate the effects that ethnic classification method has on cross-sectional quantitative analysis with child, adolescent, and adult samples—the Growing Up in New Zealand birth cohort study (for child and adult data) and the Youth2000 National Health and Wellbeing Survey Series (for adolescent data). These secondary datasets provide more comprehensive data than what could feasibly be collected via primary data collection within the course of a PhD. From these datasets, I selected the data wave for each age cohort with the most complete ethnicity data at the beginning of the data analysis phase of this PhD (April 2017). An overview of these datasets is provided below. Specific information on the analytic

samples used in the three studies of this thesis are available in their respective chapters (Chapters 4–6).

### 3.3.1 Growing Up in New Zealand

Growing Up in New Zealand is a longitudinal birth cohort study of child development in Aotearoa New Zealand (see Morton et al., 2013, for more detail). It utilises a multidisciplinary life-course approach to investigate developmental trajectories within the wider social context (e.g., familial, sociocultural, and political), with the overarching aim of informing policy for positive change. Pregnant mothers living in the Auckland and Waikato regions in the North Island of Aotearoa New Zealand, with an estimated delivery date between 25 April 2009 and 25 March 2010, were recruited via multiple strategies (e.g., through lead maternity carers, hospitals, community clinics and organisations, and the media). These regions cover over 40% of the country's total population, and were selected due to their diversity in key demographic characteristics, such as ethnicity, socioeconomic position, and urbanicity. The study cohort has 6,853 children, and is broadly nationally representative of all births in Aotearoa New Zealand between 2007 and 2010 (Morton et al., 2015). As of 2017, four major data collection waves had been conducted: antenatally, at nine months, two years, and five years. The resulting datasets include child data (collected by proxy from the child's mother or main caregiver), mother data, and partner data (typically children's biological fathers), collected via face-to-face computer-assisted personal interviews.

For this PhD, I utilised child and adult data at the Growing Up in New Zealand data collection wave that contained the most comprehensive ethnicity information for these age cohorts—namely, the five-year data collection wave for children, and the antenatal data collection wave for mothers and partners. These data waves respectively asked children (via mother proxy) and adults for (1) all the ethnic groups they identified with (i.e., total response ethnicity), and (2) the ethnicity they identified with most (i.e., self-prioritised ethnicity; see

Section 4.3.2 for more detail about these ethnicity questions). It is common practice to collect child ethnicity information via parent proxy due to their young age (Perez, 2006; Statistics New Zealand, 2005). Time-variant demographic characteristics (e.g., socioeconomic deprivation and urbanicity) were taken from the same data collection wave. Time-invariant demographic characteristics (e.g., sex) were taken from the antenatal data collection wave. Consistent with the focus of this PhD, the collated datasets were analysed in a cross-sectional (rather than longitudinal) manner.

### 3.3.2 Youth2000 Survey Series

Youth'12 data were utilised for the adolescent sample in this PhD. Youth'12 is the third survey in the multi-cohort Youth2000 National Youth Health and Wellbeing Survey Series (following Youth'01 and Youth'07; and preceding Youth'19)—a nationally representative survey series of the health and wellbeing of secondary school students in Aotearoa New Zealand (see Clark et al., 2013, for more detail). The survey series uses a multi-disciplinary ecological approach with the overarching aim of promoting healthy development and wellbeing of youth in Aotearoa New Zealand through research that is of relevance to stakeholders (Clark et al., 2013). The 2012 dataset was selected for this thesis because it was the most recent dataset when I began data analysis in 2017.[14]

Participants for Youth'12 were selected using a two-stage clustered sampling design (Clark et al., 2013).[15] First, one-third of schools (125 out of 397) in the country with Year 9

[14] In addition, the Youth'19 survey did not ask participants for their self-prioritised ethnicity, so Youth'19 data were not suitable for this thesis even when the dataset became available in 2020.

[15] This two-stage clustered sampling approach produces hierarchical, or nested, data (i.e., students nested within schools). The data were analysed using single-level techniques (rather than multi-level techniques, e.g., hierarchical linear modelling; Raudenbush & Bryk, 2002), because there was no significant school-level variation in this thesis's variables of interest (i.e., discrepancy between administratively-prioritised and self-prioritised ethnicity in Study 2, and the three mental health outcomes in Study 3; all $p > .05$).

students or above were randomly selected and invited to participate (Year 9 is the first year of secondary school in Aotearoa New Zealand; students are approximately aged 13 years). Then, within each of the 91 schools (73%) that consented to participate, 20% of students on the school roll were randomly selected and invited to participate. In smaller schools (<150 students), 30 students were randomly selected to protect confidentiality. Sampling weights were used to adjust for higher likelihood of selection in smaller schools. Of the 12,503 students who were invited to participate, 8,500 (68%) completed the survey—representing 3% of secondary school students in Aotearoa New Zealand at Year 9 or above. The anonymous online survey was administered in schools via computer-assisted self-administered interview on computer tablets. The survey was available in both English and te reo Māori (the Māori language), and had optional audio voice-over.

## 3.4 Ethical Considerations

Ethical approval for the Growing Up in New Zealand study was obtained from the Ministry of Health Northern Y Regional Ethics Committee (reference NTY/08/06/055), and ethical approval for the Youth'12 study was obtained from the University of Auckland Human Participants Ethics Committee (reference 2011/206). Data for both of these studies were collected anonymously following voluntary and informed consent (note caregivers of students under 16 years of age were able to withdraw their child from the Youth'12 study), and participants were able to skip questions they did not wish to answer. I obtained access to these datasets via data access requests to the Growing Up in New Zealand Data Access Committee, and the Adolescent Health Research Group (which oversees the Youth2000 survey series), respectively. Both data access processes involved submission of a research proposal which was reviewed by the respective teams governing the data. Research in the current thesis was conducted in accordance with the data access policies of the respective studies. This includes

submitting each of the three journal manuscripts in this thesis to the governing teams for review before they are submitted to journal outlets.

Given the ethnicity focus of this thesis, adhering to *Te Tiriti o Waitangi* obligations (see Section 2.4), as well as broader ethical principles of minimising harm (non-maleficence) and maximising benefit (beneficence), were of upmost importance throughout the research process (Hudson et al., 2010; Research Data Alliance International Indigenous Data Sovereignty Interest Group, 2019; Waitangi Tribunal, 2019). In the context of the purely secondary data analysis in this PhD, this particularly involved ensuring that research conceptualisation, data analysis, as well as interpretation and dissemination of results, align with Māori (and other ethnic minority groups') interests and benefits. To this end, this PhD had support and involvement from Māori, Pacific, and Asian cultural advisors from Growing Up in New Zealand and the Adolescent Health Research Group, many of whom also contributed to the journal manuscripts in Chapters 4 to 6 as co-authors. I also actively sought to read from, and have conversations with, Indigenous scholars to better understand their worldview(s). As discussed earlier in this chapter, I was reflexive of my positionality and potential biases throughout the whole research process, including how this influences the way I interpret data. In addition, I consciously positioned myself as an Asian tangata Tiriti and ally to Indigenous self-determination, and aligned myself with the goals of equity and social justice.

CHAPTER 4

# Study 1: Method Matters

The thesis thus far has presented the case for why ethnicity is an important variable, the need for more research on how ethnic classification method affects quantitative analysis, and how I will address this gap using a critical quantitative approach as a tangata Tiriti. This chapter, which contains the first of three peer-reviewed empirical studies, focuses on investigating the effects that researchers' ethnic classification decisions have on the outputted size of broad ethnic groupings (i.e., European, Māori, Pacific, Asian, and Other) in increasingly multi-ethnic cohorts (i.e., adults, adolescents, and children). Ethnic group size is important because it represents who is included or excluded in each ethnic group, and acts as the denominator in analyses of ethnic group differences in health and social outcomes. Ethnic group size is also important because it is implicated in the numeric visibility of a group, along with the amount of resourcing and funding each group receives.

To contextualise Study 1, I will first examine the rate of multi-ethnic identification in children, adolescents and adults, as well as each cohort's response patterns to the self-prioritisation question (including whether responses were discrepant from administrative-prioritisation, and the proportion who chose not to select a main ethnic group). Then, I will investigate how outputted ethnic group size in each age cohort varies by three ethnic classification methods: total response, administrative-prioritisation, and self-prioritisation. I focus on these three methods in this study because they each output multiple ethnicity data into a small number of broad ethnic groupings, and thus are commonly used in Aotearoa New Zealand.

The study is published in *Demographic Research*, and is reproduced in its entirety below, with minor edits to enhance consistency with the thesis. The recommended citation for this article is:

Yao, E. S., Meissel, K., Bullen, P., Atatoa Carr, P., Clark, T. C., & Morton, S. M. B. (2021). Classifying multiple ethnic identifications: Methodological effects on child, adolescent, and adult ethnic distributions. *Demographic Research*, *44*, 481–512. https://doi.org/10.4054/DemRes.2021.44.21

## 4.1 Introduction

Ethnicity is a variable widely used to measure and analyse differences between population subgroups for research and policy purposes (Balestra & Fleischer, 2018; Mays et al., 2003). Historically, boundaries between ethnic groups have been relatively clear, but with the global rise in migration and interethnic unions, ethnic boundaries are blurring, and multiple ethnic affiliations are increasingly commonplace (Aspinall, 2018b; Perez & Hirschman, 2009).[16] Many countries have adapted census and survey collection practices to support multiple ethnic responses (e.g., Aotearoa New Zealand, Australia, Canada, United Kingdom, and United States; Morning, 2008). In turn, quantitative researchers face the issue of how to classify multiple ethnic responses into a format suitable for data analysis (Callister et al., 2007; Mays et al., 2003).

There are a number of possible ethnic classification methods of varying complexity, each with associated strengths and weaknesses (Denton & Deane, 2010; Didham, 2005; Herman, 2011). However, there is scant research that comprehensively investigates the effects

[16] For brevity, "multiple" refers to "two or more" in this paper. The word "dual" will be used when specifically referring to two ethnic groups.

that different ethnic classification methods have on the outputted proportions of ethnic groups, despite the imperative of accurate ethnic group counts for robust research (Balestra & Fleischer, 2018; Mays et al., 2003). To address this gap, the current study utilises large-scale survey data from Aotearoa New Zealand to empirically explore the validity and consistency of three comparatively accessible ethnicity output methods (total response, administrative-prioritisation, and self-prioritisation) on outputted ethnic group proportions in increasingly ethnically diverse age cohorts (adults, adolescents, and children). The relevance of the study extends to other ethnically diverse countries, but especially to multi-ethnic colonised countries with Indigenous Peoples (e.g., Australia, Canada, and the United States).

### 4.1.1 Multiple Ethnic Identifications

Ethnicity is a social construct used to characterise a group of people who are perceived, by themselves and/or by others, as having shared commonalities in ancestry, history, traditions, and culture (Bhopal, 2004; Cokley, 2007). It is related to, but distinct from, the biological concept of race, which refers to groupings based on inherited physical characteristics such as skin colour, facial features, and hair texture. However, there are regional differences in the usage of these terms—for example, in the United States, the term "race" is often used instead of "ethnicity" (Morning, 2008).[17] Traditionally, ethnicity has been considered a time-invariant construct with mutually exclusive groups. However, due to increasing transnational mobility, interethnic unions, multi-ethnicity, and self-ascribed understandings of identity, ethnicity is now better understood as a complex, dynamic construct (Aspinall, 2018b; Bhopal, 2004; Morning, 2008).

[17] The term "ethnicity", rather than "race", will be used throughout this paper except when it pertains directly to the U.S. context. In these instances, the use of race is synonymous with ethnicity.

Globally, a steadily growing proportion of the population report identification with more than one ethnic group (Aspinall, 2018b; Balestra & Fleischer, 2018). In Aotearoa New Zealand, 2018 census data showed that 11.4% of the population identified with two or more broad ethnic groupings, up from 9.0% in 2001, 10.4% in 2006, and 11.2% in 2008 (Statistics New Zealand, 2014b, 2020a). In the United States, 9 million individuals (2.9%) reported two or more races in the 2010 Census, up from 6.8 million (2.4%) in the 2000 Census (N. A. Jones & Bullock, 2012). The prevalence of multiple ethnic identifications is usually higher in younger age groups. For example, in Aotearoa New Zealand's 2018 census, 23.5% of children (aged 0–14 years) were reported as belonging to more than one ethnic grouping, compared to just 2.8% of those aged 65 years and over (Statistics New Zealand, 2020a). Based on the youthfulness of the multi-ethnic population, as well as increasing societal acceptance of multiple ethnic identifications, the size of this group is projected to continue to grow (Aspinall, 2018b; Perez & Hirschman, 2009).

Recognising the increasing global trend in multiple ethnic identifications, the United Nations (2008) updated the *Principles and Recommendations for Population and Housing Censuses* to require that "respondents have the option of indicating multiple ethnic affiliations" (p. 139) in census forms. Sample surveys also increasingly allow for multiple ethnic responses. There is no single internationally recommended standard for the collection of multiple ethnicity data, as this is dependent upon each country's ethno-racial history, current ethnic composition, and socio-political context (Balestra & Fleischer, 2018; United Nations, 2008). However, common data collection methods for those identifying with multiple ethnicities can be classified into four categories (Aspinall, 2018b; Morning, 2008): (1) check-all-that-apply on a predetermined list including an open-ended "other" response option (e.g., Aotearoa New Zealand, Canada, the United States), (2) specific combinations of ethnic groups including an open-ended "other" response option (e.g., England, Wales), (3) a generic "mixed" option (e.g.,

Barbados, Jamaica, St Lucia), and (4) an open-ended free-text box (e.g., Northern Ireland, Scotland). Irrespective of the format of data collection, researchers working with multiple ethnicity data face the identical issue of how to classify these responses and ensure that statistical output provides an appropriate interpretation of the data collected.

### 4.1.2 Ethnic Classification Methods

Ethnicity data are commonly used to research and monitor ethnic inequities in health and social outcomes (Balestra & Fleischer, 2018; Mays et al., 2003). In addition, they are used to target and evaluate policies, funding, services, and interventions aimed at reducing ethnic inequities, and thus are especially important in colonised countries with legal and moral obligations to Indigenous Peoples. As a result, ethical, accurate, and consistent classification of ethnicity data is paramount to affirmative action and antidiscrimination agendas in research and policy. There is some empirical research which suggests that the choice of ethnic classification method can affect the interpretation of ethnic differences in health outcomes (e.g., Boven et al., 2020; Callister et al., 2007; Mays et al., 2003; Rutkowski et al., 2017). Ethnic classification, as used in this paper, refers to the coding or output of multiple ethnic responses into a format suitable for data presentation (e.g., in tables) or further statistical analyses (e.g., as an independent variable).

Methods for classifying multiple ethnicity data have been described in various sources, including textbooks (e.g., Denton & Deane, 2010; Subramanian, 2009), academic journals (Atatoa Carr et al., 2017; Herman, 2011), and reports (Didham, 2005; S. M. Lee, 2001). The literature in this area is primarily from Aotearoa New Zealand and the United States, and tends to draw key principles from official governmental (Statistics New Zealand, 2005) and federal (Office of Management and Budget [OMB], 2000) documents on race/ethnicity statistics. Table

4.1 provides an overview of common ethnic classification methods.[18] Although the terminology of ethnic classification methods differs between the two countries and the finer details of implementation diverge (e.g., aggregation of racial/ethnic groups), two broad approaches can be identified: (1) methods that *retain* multiple ethnicity data, and (2) methods that *reduce* multiple ethnicity data. The method implemented is often not made transparent in academic journal publications, so patterns of usage are not readily known.

Ethnic classification methods can be evaluated based on their ethical and statistical appropriateness, as well as their general usability (Mays et al., 2003; OMB, 2000). Because ethnicity is a measure of self-affiliation, it is both ethically and analytically important that the classification method is as congruent as possible with individuals' initial responses. However, this may need to be balanced against fundamental criteria for robust statistical analysis. For instance, adequate subgroup sample sizes are needed to retain explanatory power; and subgroup membership needs to be relatively consistent across time and contexts, particularly for longitudinal analyses, or where numerator and denominator data are collected at different times or contexts. In addition, an ideal classification method should be fairly accessible for researchers to implement and easily understood by the research audience (Mays et al., 2003; OMB, 2000).

[18] Less common methods include fractional assignment, which allocates a weighting to each selected ethnic group so they sum to 1 (e.g., 0.5 weighting to each of the two ethnic groups selected by a dual-ethnic participant; OMB, 2000), and the National Health Interview Survey (NHIS) regression method, which uses individual and contextual characteristics to predict a participants' "main" race (Ingram et al., 2003). These methods were not considered appropriate for the Aotearoa New Zealand context because they contradict the country's official definition of ethnicity (i.e., they relegate a multi-ethnic participant into fractions, and deny ethnic self-identification, respectively; Statistics New Zealand, 2005).

**Table 4.1**

*Overview of Common Ethnic Classification Methods*

| Ethnic classification method | Alternative names | Description |
|---|---|---|
| *Retention methods* | | |
| Sole/combination grouping | • Single/combination grouping<br>• Detailed ethnicity<br>• Complex ethnicity | Participants are counted *once* according to the ethnic group or combination of ethnic groups they report (e.g., a person who identified as both Māori and European is classified as Māori/European). |
| Total response grouping | • All-inclusive method | Participants are counted in *each* of the ethnic groups they report. Total counts can exceed total participants because participants who report more than one ethnic group are counted more than once (e.g., a person who identified as both Māori and European is classified in both the Māori and European ethnic groups). |
| *Reduction methods* | | |
| Administrative-prioritisation | • Prioritisation<br>• External prioritisation<br>• Deterministic whole assignment | Participants who report multiple ethnic groups are assigned to a *single* category based on a *predetermined hierarchy*. The hierarchy used in Aotearoa New Zealand is Māori > Pacific > Asian > Other > European (e.g., a person who identified as both Māori and European is classified as Māori, as Māori is above European in the hierarchy). |
| Self-prioritisation | • "Main" ethnic group<br>• "Best" ethnic group<br>• "Primary" ethnic group | Participants who report multiple ethnic groups are asked in a follow-up question to select *one* ethnic group that they identify with *most* (e.g., a person who identified as both Māori and European, and then identified European as their main ethnic group, is classified as European). |

#### *4.1.2.1 Retention of Multiple Ethnicities*

Classification methods that retain multiple ethnicity data preserve multiple ethnic respondent affiliation. However, these methods tend to pose more statistical challenges. *Sole/combination grouping* is the classification method that outputs multiple ethnic identifications into a format most similar to original responses (OMB, 2000; Statistics New Zealand, 2005). Sole/combination grouping collapses populations into single or combinations of broad ethnic groups. For example, the broad ethnic groups in Aotearoa New Zealand are European; Māori (Indigenous Peoples of Aotearoa New Zealand); Pacific Peoples; Asian; Middle Eastern, Latin American, or African (MELAA); and Other (a residual category). Therefore, sole/combination grouping would comprise European only, Māori only, Pacific only, Māori/European, Pacific/European, Māori/Pacific/European, Māori/Pacific/Asian, and so on. This method enables more nuanced insight into ethnic patterns compared to other ethnic classification methods that rely on broad ethnic groupings. For instance, those who affiliate with multiple ethnic groups across these broad categories may have different characteristics and outcomes than those who affiliate solely within one ethnic group (Callister et al., 2007; Didham, 2005). However, detailed combinations can result in an unmanageably large number of subgroup permutations, and when ethnicities are classified in this way, there tends to be more instability in ethnic group membership over time (Didham, 2005). Further, some combination groupings may have sample sizes that are too small for meaningful statistical analyses. A commonly practiced solution is to aggregate combinations with small cell counts into broader categories as appropriate (e.g., “other dual combinations”, “three or more ethnic groups”, or sometimes a single “mixed” category; Statistics New Zealand, 2005), but the additional heterogeneity within these residual categories compromises the level of nuance that can be achieved.

An alternative common classification method which retains multiple ethnicity data is *total response grouping* (OMB, 2000; Statistics New Zealand, 2005). Rather than specifying the exact combination of responses, this approach assigns individuals to all the ethnic groups they identify with, whether alone or in combination. For example, the European grouping includes those who identify as European only, and those who identify within the European group as well as other ethnic group(s). This means that individuals who identify with multiple groups are counted multiple times, as each ethnic group is treated as a separate binary *yes/no* variable. The main advantages of this approach are its clear indication of the number of respondents who identify with each ethnic group, as well as its ability to classify ethnic groups in more specificity (e.g., ethnicities within broad ethnic groupings). However, mono- and multi-ethnic responses cannot be differentiated using this approach, the outputted categories are not mutually exclusive, and the sum of ethnic group counts is greater than the number of participants. This can introduce complexities for some statistical analyses and make results less intuitive to readers. Additionally, numerically dominant groups can subsume smaller-sized ethnic groups, which challenges the ability to appropriately measure, monitor, and address inequities (Department of Statistics, 1993).

#### *4.1.2.2 Reduction of Multiple Ethnicities*

In contrast, classification methods which reduce multiple ethnic affiliations into a single ethnic grouping tend to be more attractive for some statistical procedures, but may be ethically problematic and subject to interpretative inaccuracy, because such methods do not utilise all the information provided by respondents. There are two main approaches in this category. The first approach, known as *administrative-prioritisation* in Aotearoa New Zealand and *deterministic whole assignment* in the United States, follows a predetermined set of rules to allocate multiple responses into a single ethnic group (OMB, 2000; Statistics New Zealand, 2004). In Aotearoa New Zealand, the standard administrative-prioritisation hierarchy

prioritises Māori responses, followed by Pacific, Asian, ethnic groups other than European, and finally, European responses (Department of Statistics, 1993). The rationale behind this hierarchy is to first prioritise Indigenous Peoples in recognition of Indigenous rights under Aotearoa New Zealand's founding document, *Te Tiriti o Waitangi* (*The Treaty of Waitangi*),[19] and then prioritise the ethnic groups of policy relevance and those which are smaller in size. In the United States, deterministic whole allocation is usually based on the relative size of racial groups (e.g., "Smallest Group", "Largest Group", or "Largest Group Other Than White"). These were first outlined by the OMB (2000) as "bridging methods" to ensure a smooth transition when multiple race data were first collected in the 2000 Census. The major advantage of a predetermined algorithm is its ability to classify multiple ethnic responses into a small number of mutually exclusive categories without needing to seek additional information from respondents. However, external allocation to a single ethnic group may be ethically contentious because it is dismissive of the self-affiliation of multi-ethnic participants. Existing evidence from survey research in Aotearoa New Zealand that asked multi-ethnic respondents to select a "main" ethnic group indicates that the discrepancy between the two methods is around 40% in both adolescents (Kukutai & Callister, 2009) and adults (Atatoa Carr et al., 2017). External allocation can also be problematic because it skews ethnic distributions by understating the size of ethnic groups that are not in first position in the algorithm.

The alternative approach is *self-prioritisation*, where multi-ethnic respondents are asked in a follow-up question to select their "main" ethnic group, thereby taking strength of self-affiliation into account. This approach is more commonly adopted in sample surveys (e.g., Atatoa Carr et al., 2017; D. R. Harris & Sim, 2002), and is rarely observed in population

[19] *Te Tiriti o Waitangi* is a constitutional document stipulating an agreement between Indigenous Māori and the British Crown, whereby Māori are guaranteed the rights to partnership, participation, and protection.

censuses. An exploratory study in Aotearoa New Zealand found that, in a sample of 641 multi-ethnic early adolescents, nearly three-quarters were willing and able to select a main ethnic group when required (Kukutai & Callister, 2009). This suggests that self-prioritisation can be a viable method for many people, but the issue of how to categorise those who do not self-prioritise an ethnicity remains. Regardless of whether multi-ethnic participants are willing to select a main ethnic group, some proponents argue that it is unethical to ask them to do so, because it forces them to elevate one ethnicity over another and conveys that multi-ethnic identity is unvalued (Sanchez, 2010; Townsend et al., 2009). Furthermore, while all ethnic identification is context-specific, self-prioritised ethnicity may be particularly susceptible to socio-contextual influences such as racism and stigma (Herman, 2004). As a result, responses may be inauthentic, or change between contexts such as home and school (D. R. Harris & Sim, 2002).

## 4.2 The Current Study

The steadily growing size of the global multi-ethnic population, in conjunction with the importance of accurate and consistent ethnic group counts for research and policy, pose the pressing and complex question of how best to classify multiple ethnic responses for statistical analyses. As discussed in the previous section, the literature outlines a number of possible strategies, each with advantages and disadvantages. However, the field lacks empirical research that comprehensively compares these methods. The current study utilises large-scale survey data from Aotearoa New Zealand to address this gap. The study will be of interest to other colonised countries which have a similar historical background and face analogous contemporary issues related to ethnicity statistics.

### 4.2.1 Aotearoa New Zealand Context

Aotearoa New Zealand is a valuable context for the current study because of its ethnic diversity and relatively large Indigenous Māori population (17%; Statistics New Zealand, 2020a). European colonisers began arriving in this country from the early 1800s and outnumbered Māori by 1858. A small Chinese population also settled during the Gold Rush in the 1860s. Significant waves of migration have followed over the last 150 years, including from the Pacific (particularly Sāmoa, Tonga, and Fiji) following World War II to support labour in secondary industries and, after a loosening of immigration policy, modern migration of Asian Peoples from the 1970s. More recently, migrants and refugees from the Middle East, Latin America, and Africa have settled in the country (Callister et al., 2007; Khawaja et al., 2000). Interethnic marriage in the country has occurred since early European colonisation (Callister et al., 2007).

Aotearoa New Zealand currently has a population of around 5 million people. Its five-yearly census collected multiple ethnicity data from 1991 onwards (Khawaja et al., 2000). In the 2018 Census, 11% of the population reported multiple ethnicities (Statistics New Zealand, 2020a). Multiple affiliations were more common in Māori and Pacific Peoples, and among younger age groups. The main ethnic groups in Aotearoa New Zealand, in descending proportional size of total responses enumerated in the 2018 Census, are: European (70%), Māori (17%), Asian (15%), and Pacific Peoples (8%; Statistics New Zealand, 2020a). It should be noted that, except for Māori, these ethnic groups are broad ethnic groupings utilised for statistical output only, and that there is considerable heterogeneity within each group (Statistics New Zealand, 2005).

Statistics New Zealand's (2005) *Statistical Standard for Ethnicity* provides official guidelines on the collection and output of ethnicity data based on Statistics New Zealand's (2004) *Report of the Review of the Measurement of Ethnicity*. Regarding ethnicity output, the

review panel assessed the relative merits and limitations of possible ethnic classification methods, and recommended that sole/combination and total response groupings be used as standard output in place of the previously endorsed process of administrative-prioritisation. The panel did not consider self-prioritisation an appropriate alternative because, like administrative-prioritisation, it "contradicts the concept and definition of ethnicity [whereby people can belong to more than one ethnic group]" (Statistics New Zealand, 2004, p. 13). Although the panel's recommendations were adopted in the country's official statistical standard for ethnicity, administrative-prioritisation continues to be routinely used in health and education research, particularly when addressing inequity (Education Counts, 2014; Ministry of Health, 2017).

### 4.2.2 Research Aims

The overarching aim of the current study is to comprehensively examine three ethnic classification methods (total response, administrative-prioritisation, and self-prioritisation) in order to ascertain whether these have a differential effect on outputted ethnic group proportions in age cohorts of differing levels of ethnic diversity (children, adolescents, and adults) in Aotearoa New Zealand. These classification methods were selected because they are the most accessible methods that can output multiple ethnic responses as a small number of broad ethnic groupings. The specific research questions investigated in this study are as follows:

1. What proportion of multi-ethnic participants in each age group selected a main ethnic group when asked?
2. What is the rate of alignment between administratively-prioritised and self-prioritised ethnicity for multi-ethnic participants in each age group?
3. How do overall ethnic group proportions (i.e., including mono-ethnic participants) in each age group differ by total response, administratively-prioritised, and self-prioritised output?

## 4.3 Method

### 4.3.1 Data Sources

The current study utilises data from two large studies in Aotearoa New Zealand: Growing Up in New Zealand and the Youth2000 National Youth Health and Wellbeing Survey. Growing Up in New Zealand is a longitudinal birth cohort study which recruited pregnant mothers in the ethnically diverse regions of Auckland and Waikato in the North Island of Aotearoa New Zealand with an estimated due date in 2009 or 2010 (see Morton et al., 2013). The cohort is broadly generalisable in key demographic characteristics to the Aotearoa New Zealand birth population (Morton et al., 2015). Current partners of the mothers were also recruited independently through contact information provided by mothers.

Mother and partner data from the Growing Up in New Zealand antenatal data collection wave formed the adult sample in the present study. This data wave was selected because it collected the most detailed ethnicity information from adults. Participants with non-missing total response ethnicity (99.9%) were utilised as the analytic sample ($N$ = 11,210). There were more mothers ($N$ = 6,814; 61%) than partners ($N$ = 4,396; 39%) in this sample. The median age of partners (33 years) was slightly higher than the median age of mothers (31 years). Nearly all partners (99%) were biological fathers of the children in the cohort.

Detailed ethnicity data for children were collected by proxy from the child's mother (or main caregiver) at the 54-month data collection wave. Participants with non-missing total response ethnicity (99.9%) were utilised as the child sample in the present study ($N$ = 6,149). Ethical approval for Growing Up in New Zealand was obtained from the Ministry of Health Northern Y Regional Ethics Committee (reference NTY/08/06/055). All Growing Up in New Zealand data utilised in the present study were collected by face-to-face computer-assisted personal interviews.

The adolescent sample for the current study was drawn from the Youth'12 dataset. This dataset was collected in 2012 as part of the nationally representative cross-sectional Youth2000 survey series on the health and wellbeing of secondary school students in Aotearoa New Zealand (aged 12–18 years; see Clark et al., 2013). Data were collected in schools via a multimedia computer-administered survey on computer tablets. Participants with non-missing total response ethnicity (99.6%) were utilised as the analytic sample ($N$ = 8,464). Ethical approval for Youth'12 was obtained from The University of Auckland Human Participants Ethics Committee (reference 2011/206).

### 4.3.2 Measures

The dataset for each age group had two survey items that were used for the purposes of this study (note that all child data were collected by proxy from mothers). The first item was *total response ethnicity*, where participants were asked to select *all* the ethnic groups they identified with. The second item was *self-prioritised ethnicity*, where participants who selected more than one ethnic group were asked to select the ethnicity that they identified with *most*. Growing Up in New Zealand allowed participants to select two main ethnic groups if necessary and also provided a "don't know" option, whereas Youth'12 provided an option of "I can't choose only one ethnic group". The Growing Up in New Zealand survey also allowed participants to refuse response. Table 4.2 shows the questions and response options for each age group in more detail.

Responses to both ethnicity items were collected at a detailed level of ethnicity (see Table 4.2), but were aggregated into five broad groupings for the purposes of this study:

**Table 4.2**

*Ethnicity Questions and Response Options by Age Group*

| | Children[a] | Adolescents | Adults |
|---|---|---|---|
| *Dataset* | | | |
| Study | Growing Up in NZ | Youth2000 survey series | Growing Up in NZ |
| Data collection wave | 54 months | Youth'12 | Antenatal |
| *Total response ethnicity* | | | |
| Question | "Which ethnic group or groups does {NAME} belong to? (Tick all that apply—at least one)" | "Which ethnic group do you belong to? (you may choose as many as you need)" | "Which ethnic group OR GROUPS do you belong to? (Choose the answer or answers that apply to you)" |
| Response type | Multiple response | Multiple response | Multiple response |
| No. of response options | 33 Level 3 categories[b] (including open-ended "Other—please specify") | 24 Level 2 categories[b] (including closed-ended "Other") | 33 Level 3 categories[b] (including open-ended "Other—please specify") |
| *Self-prioritised ethnicity* | | | |
| Question | "Which is the MAIN ethnic group that {NAME} identifies with?" | "Which is your main ethnic group (the one you identify with most)?" | "Which is your main ethnic group that is the one you identify with most?" |
| Response type | Maximum of 2 responses | Single response | Maximum of 2 responses |
| No. of response options | 33 Level 3 categories[b] (including open-ended "Other—please specify") + "don't know" + refused | 23 Level 2 categories[b] (including closed-ended "Other") + "I can't choose only one ethnic group" | 33 Level 3 categories[b] (including open-ended "Other—please specify") + "don't know" + refused |

[a]Child responses were collected by proxy from the child's mother.

[b]Levels are broadly based on Statistics New Zealand's (2005) four-level hierarchical ethnic classification system of increasing specificity in ethnic groups. Level 1 has six categories (European, Māori, Pacific, Asian, MELAA, and Other) and is used solely for output. Excluding residual categories, Level 2 has 21 categories (e.g., Sāmoan, Cook Islands Māori, Chinese, Indian), Level 3 has 36 categories (e.g., Italian, German, Filipino, Cambodian), and Level 4 has 180 categories (e.g., Dalmatian, Macedonian, Papua New Guinean, Burmese, Malaysian).

European,[20] Māori, Pacific Peoples, Asian, and Other. Other was a residual category that included MELAA (which had a relatively low proportion of responses) and New Zealander (which was not a specified response option but was coded from open-ended responses where available). In this study, multi-ethnic participants were defined as those who selected more than one ethnicity at this broad grouping level. Data for the other ethnic classification methods used in this study were derived from aggregated total response data. Specifically, *sole/combination ethnicity* was coded from total response ethnicity into 13 groups (see Table 4.3 for a group list). Finer groupings were not possible due to small cell counts ($n < 20$). *Administratively-prioritised ethnicity* was derived from total response ethnicity using the hierarchy specified by Aotearoa New Zealand's Department of Statistics (1993): Māori responses were prioritised, followed by Pacific, Asian, Other, and finally European responses.

A binary variable was then created to indicate instances where there was a discrepancy between administratively-prioritised ethnicity and self-prioritised ethnicity for multi-ethnic participants. In this study, a discrepancy was operationalised as a case where administratively-prioritised and self-prioritised ethnicity did not match at the broad level of ethnic grouping used, or where participants did not select a main ethnic group (i.e., selected two main ethnic groups, "don't know", "I can't choose only one ethnic group", or refused response). In both scenarios, the result of administrative-prioritisation differs from the response given by the participant. These two scenarios were aggregated to mitigate comparability concerns due to differing survey response options for adolescents.

[20] The majority of participants in the European grouping identified as New Zealand European (>85%). The remaining participants identified as other European ethnicities (e.g., Australian, British, Dutch, and Irish).

### 4.3.3 Data Analysis

For each research question, descriptive statistics for each age group were calculated. Sample size was used as the denominator to calculate ethnic group percentages for all ethnic classification methods. Then a series of two-sample *z*-tests for proportions without continuity correction were conducted to test for age group differences in multi-ethnic participants' selection of a main ethnic group (Research Question 1), and age group differences in discrepancies between administratively-prioritised and self-prioritised ethnicity (Research Question 2). Specifically, three pairwise age-group comparisons were conducted to answer each research question: children versus adolescents, children versus adults, and adolescents versus adults. Due to the sensitivity of the *z*-test to large sample sizes and the inflated family-wise Type I error rate arising from multiple comparisons (Field et al., 2012), an alpha level of $\alpha = .001$ was adopted to indicate statistical significance.[21]

To compare how overall ethnic group proportions (i.e., including mono-ethnic participants) differed by ethnic classification method for each age group (Research Question 3), three one-sample *z*-tests for proportions without continuity correction were conducted for each broad ethnic grouping within each age group. Participants who selected multiple ethnic groups but did not indicate a self-prioritised ethnicity were excluded to maximise comparability across age groups (adolescents had more restrictive self-prioritisation response options than children and adults; see Table 4.2). Total response was used as the reference population for comparisons with the two prioritisation methods, and self-prioritisation was used as the reference population for comparison with administrative-prioritisation. One-sample *z*-tests

[21] Statisticians increasingly deem dichotomisation of statistical significance problematic (see Wasserstein et al., 2019, for a discussion), but it was included in this study because it was considered to be a pragmatic indicator of whether the differences observed were likely due to chance or not. Precise *p*-values are available in Appendix A, and statistical testing was complemented by effect size calculations.

were used because the proportions being compared were from the same population; two-sample *z*-tests were not appropriate as they assume independence between samples. An alpha level of $\alpha = .001$ was again used to indicate statistical significance.

For all three research questions, non-directional Cohen's (1988) *h* effect size ($|h|$), which calculates the absolute difference in the arcsine transformations of two proportions, was used to measure the magnitude of difference between proportions. This effect size measure is appropriate for both independent and matched samples (Cohen, 1988). Cohen's interpretation of $|h|$ was adopted: 0.20 was considered a small effect, 0.50 was considered a medium effect, and 0.80 was considered a large effect. Sampling weights were used in all analysis of adolescent data to adjust for intentional oversampling of students in smaller schools (Clark et al., 2013). Sampling weights were not required for child and adult data as the Growing Up in New Zealand study was broadly representative of the Aotearoa New Zealand birth cohort (Morton et al., 2015). All analyses were conducted using R version 3.6.1 (R Core Team, 2019).

## 4.4 Results

### 4.4.1 Descriptive Statistics

Participants' ethnic identification (which, for children, was selected via mother proxy) are presented by age group using sole/combination grouping in Table 4.3. This ethnic classification method displays information about the number of broad ethnic groups each participant identified with, as well as the specific ethnic group or groups at the level of aggregation used. Table 4.3 shows that although the majority of participants in each age group identified with one broad ethnic grouping, there was higher ethnic diversity among young people—children (by proxy) and adolescents were more than twice as likely to identify with more than one broad ethnic grouping compared to adults. Of those who identified with more than one broad ethnic grouping, most identified with two broad ethnic groups; only a relatively

**Table 4.3**

*Sole/Combination Ethnic Identification by Age Group*

| | Children[a] | | Adolescents | | Adults | |
|---|---|---|---|---|---|---|
| | *n* | Col. % | *n* | Col. % | *n* | Col. % |
| *N* | 6,149 | | 8,464 | | 11,210 | |
| *1 ethnic group* | | | | | | |
| European only | 2,221 | 36 | 3,989 | 47 | 5,639 | 50 |
| Māori only | 168 | 3 | 295 | 3 | 671 | 6 |
| Pacific only | 531 | 9 | 546 | 6 | 1,256 | 11 |
| Asian only | 604 | 10 | 760 | 9 | 1,543 | 14 |
| Other only | 266 | 4 | 162 | 2 | 343 | 3 |
| **Total** | **3,790** | **62** | **5,752** | **68** | **9,452** | **84** |
| *2 ethnic groups* | | | | | | |
| Māori/European | 725 | 12 | 986 | 12 | 974 | 9 |
| Pacific/European | 195 | 3 | 393 | 5 | 217 | 2 |
| Asian/European | 161 | 3 | 235 | 3 | 61 | 1 |
| Other/European | 374 | 6 | 332 | 4 | 124 | 1 |
| Māori/Pacific | 144 | 2 | 78 | 1 | 109 | 1 |
| Other 2 ethnic groups | 247 | 4 | 238 | 3 | 111 | 1 |
| **Total** | **1,846** | **30** | **2,262** | **27** | **1,596** | **14** |
| *3+ ethnic groups* | | | | | | |
| Māori/Pacific/European | 178 | 3 | 126 | 1 | 86 | 1 |
| Other 3+ ethnic groups | 335 | 5 | 324 | 4 | 76 | 1 |
| **Total** | **513** | **8** | **450** | **5** | **162** | **1** |

[a]Child responses were collected by proxy from the child's mother.

small proportion of each age group identified with three or more broad ethnic groupings. Within each age group, those who identified Māori as one of their ethnic groups were substantially more likely to identify with more than one broad ethnic grouping (children = 89%; adolescents = 83%; adults = 65%).

Sole/combination ethnicity information in Table 4.3 also shows that, as the dominant culture, "European only" was by far the most common identified ethnicity in each age group. Of the single broad ethnic groupings, this was followed by "Asian only", "Pacific only", and, for adolescents and adults, "Māori only". Māori/European was the most common combination

grouping in each age group and the proportion was larger than "Māori only" in each instance. The prevalence of other combination groupings was relatively low, and some combinations had to be aggregated due to small cell counts ($n < 20$). Combinations that included the European grouping (e.g., Pacific/European, Māori/Pacific/European) tended to be more common and hence could be presented in a disaggregated manner; the most common combination without the European grouping was Māori/Pacific.

### 4.4.2 Multi-Ethnic Participants' Selection of a Main Ethnic Group

Participants who identified with more than one ethnic group were asked in a follow-up question to select a main ethnic group. Our first research question investigated the proportion of multi-ethnic participants in each age group who did so. Due to survey differences, children (by mother proxy) and adults had the option of selecting up to two main ethnic groups, "don't know", or refuse response; whereas adolescents had the option of selecting either one main ethnic group or "I can't choose only one ethnic group". Responses for each age group were aggregated into two categories: those who selected one main ethnic group, and those who did not select one main ethnic group (i.e., selected two main ethnic groups, "don't know", "I can't choose", or refused response). As shown in Table 4.4, a clear majority in each age group selected one main ethnic group.

Pairwise age group comparisons conducted using two-sample *z*-tests (see Supplementary Table A.1 in Appendix A) showed that the proportion of adolescents who selected one main ethnic group was higher than children and adults (both $p < .001$). The magnitude of effect was medium in size ($|h| = 0.51$ and 0.60, respectively). However, caution is needed when interpreting these results, because alternative response options for adolescents were more restrictive, and child responses were via mother proxy. There were no differences in the pairwise comparisons of proportions of children and adults who selected one main

**Table 4.4**

*Dual- and Multi-Ethnic Participants' Self-Selection of a Main Ethnicity by Age Group*

| | Children[a] | | Adolescents | | Adults | |
|---|---|---|---|---|---|---|
| | *n* | Col. % | *n* | Col. % | *n* | Col. % |
| *N* | 2,359 | | 2,712 | | 1,758 | |
| *Selected* one *main ethnic group* | | | | | | |
| Matched admin-prioritisation | 732 | 31 | 1,118 | 41 | 677 | 39 |
| Did *not* match admin-prioritisation | 1,082 | 46 | 1,483 | 55 | 733 | 42 |
| **Total** | **1,814** | **77** | **2,601** | **96** | **1,410** | **80** |
| *Did* not *select one main ethnic group*[b] | | | | | | |
| Two main ethnic groups selected | 466 | 20 | - | - | S | 19 |
| "Don't know" or refused response | 79 | 3 | - | - | <10 | <1 |
| "I can't choose only one ethnic group" | - | - | 111 | 4 | - | - |
| **Total** | **545** | **23** | **111** | **4** | **348** | **20** |

*Note.* Cell counts less than 10 are suppressed as "<10". Secondary suppression (S) was applied to the next smallest cell in the column so that the suppressed cell cannot be recalculated.

[a]Child responses were collected by proxy from the child's mother.

[b]Children (by proxy) and adults could select a maximum of *two* main ethnic groups, "don't know", or refuse response. Adolescents could either select *one* main ethnic group or the "I can't choose only one ethnic group" option.

ethnic group ($p > .001$). The majority of children and adults who did not select one main ethnic group selected two main ethnic groups, rather than "don't know" or refusing to respond.

A more detailed table which disaggregates the selection of a main ethnic group by age group and ethnic combination is available in Supplementary Table A.2 in Appendix A. Examination of responses by ethnic combination was not a main focus of the study. However, it should be noted that there was considerable heterogeneity in responses between ethnic combinations.

### 4.4.3 Alignment Between Administrative-Prioritisation and Self-Prioritisation

Administrative-prioritisation is another method that can be used to categorise multi-ethnic identifications into a small number of mutually exclusive ethnic groupings. Our second research question investigated the rate of alignment between administratively-prioritised and

self-prioritised ethnicity for multi-ethnic participants in each age group. Frequencies of instances where administratively-prioritised ethnicity was different from participants' self-selected main ethnic group, as well as frequencies of participants who did not self-select a main ethnic group, can be found in Table 4.4. Aggregation of these respective rows result in the overall rate of discrepancy between the two prioritisation methods as operationalised in this study. In descending order, the overall discrepancy rates were: 69% for children, 61% for adults, and 59% for adolescents.

Pairwise age group comparisons conducted using two-sample *z*-tests (see Supplementary Table A.3 in Appendix A) showed that the discrepancy rate in children was higher than in both adolescents and adults (both $p < .001$). The magnitude of the effect was small in size ($|h| = 0.21$ and 0.16, respectively). There was no difference in the discrepancy rates between adolescents and adults ($p > .001$) despite the more restrictive alternative response options for adolescents.

Discrepancy rates by ethnic combination can be derived from Supplementary Table A.2 in Appendix A. Once again, there was considerable within-group heterogeneity. For example, the discrepancy rate was higher in Asian/European children (92%), Asian/European adults (79%), and Other/European adolescents (85%) when compared to their Māori/European counterparts (children = 76%; adolescents = 65%; adults = 63%).

### 4.4.4 Effect of Ethnic Classification Method on Ethnic Group Proportions

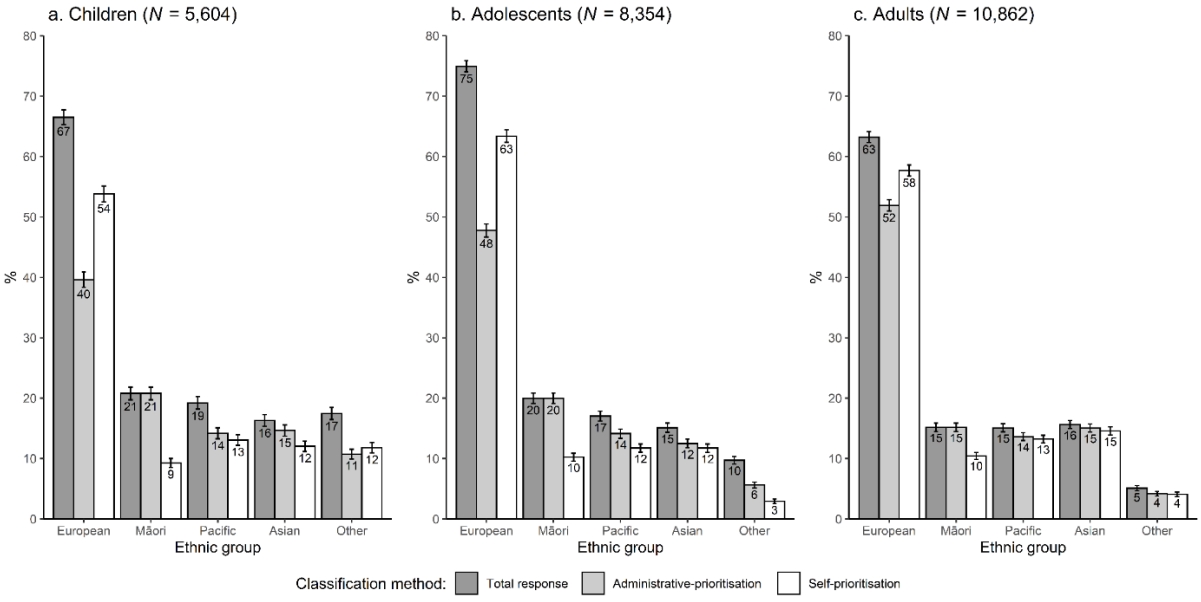
Our final research question investigated how overall ethnic group proportions (i.e., including mono-ethnic participants) in each age group differed by ethnic classification method (i.e., total response, administrative-prioritisation, and self-prioritisation). Multi-ethnic participants who did not select a main ethnic group (see final row in Table 4.4) were excluded from analyses to mitigate potential differences across age groups arising from the more restrictive self-prioritisation response options for adolescents. Note that the sum of total

response ethnicity (children = 7,862; adolescents = 11,425; adults = 12,404) was larger than the sample size by around 40% in young people and 14% in adults.

As Figure 4.1 shows, European and Māori proportions were most affected by ethnic classification method, and the effect of ethnic classification method on ethnic group proportions was stronger in young people than adults. Pairwise comparisons of the effect of ethnic classification method on ethnic group proportions for each age group were conducted using one-sample *z*-tests. Results (see Supplementary Tables A.4 to A.6 in Appendix A) showed that in each age group, the proportion of Europeans as outputted by total response was higher than self-prioritisation (all $p < .001$), which in turn outputted a higher proportion of Europeans than administrative-prioritisation (all $p < .001$). The proportion of Māori outputted by total response and administrative-prioritisation, which were identical as Māori has first position on the prioritisation hierarchy, were also higher than self-prioritisation of Māori in each age group

**Figure 4.1**

*Ethnic Group Proportions by Ethnic Classification Method and Age Group*




*Note.* Participants who did not select a main ethnic group were excluded. Child responses were collected by proxy from the child's mother. Error bars show 95% confidence intervals.

(all $p < .001$). For children and adolescents, self-prioritisation halved the outputted Māori proportion. In general, the effect of ethnic classification method on outputted European and Māori proportions was around twice as large in magnitude in children and adolescents ($|h|$ between 0.25 and 0.57) compared to adults ($|h|$ between 0.11 and 0.23).

The effect of ethnic classification method was much smaller in the other three broad ethnic groupings (i.e., Pacific, Asian, and Other). For children and adolescents (see Supplementary Tables A.4 and A.5 in Appendix A), total response also outputted higher ethnic group proportions than administrative-prioritisation and self-prioritisation (all $p < .001$), and administrative-prioritisation outputted higher ethnic group proportions than self-prioritisation in some cases ($p < .001$ for Asian children, Pacific adolescents, and adolescents of other ethnicities). However, the magnitude of effect was generally trivial ($|h| < 0.20$). In adults (see Supplementary Table A.6 in Appendix A), total response outputted higher Pacific and Other proportions than administrative-prioritisation and self-prioritisation (all $p < .001$), but the differences observed (<2%) were very trivial (all $|h| < 0.06$). No differences were observed in the other comparisons conducted for adults (all $p > .001$).

## 4.5 Discussion

The current study examined large-scale survey data of children, adolescents, and adults from Aotearoa New Zealand to investigate the validity and consistency of three relatively accessible ethnic classification methods: total response, administrative-prioritisation, and self-prioritisation. Consistent with national and global trends (e.g., Aspinall, 2018b; Statistics New Zealand, 2014b), there were higher rates of multi-ethnic identification among the younger age groups in the datasets utilised. Analysis of the data indicated that the majority of multi-ethnic participants (by mother proxy for children) selected a main ethnic group when required, and administratively-prioritised ethnicity was discrepant from participant responses over 60% of

the time. At the overall sample level (i.e., including mono-ethnic participants), the three ethnic classification methods produced within-group differences in ethnic group proportions, particularly in subgroups with higher rates of multi-ethnicity. Some age variation was observed. The results are discussed in more detail in the following sections.

### 4.5.1 Validity of Ethnicity Prioritisation Methods

The first research question investigated the proportion of multi-ethnic participants in each age group who selected a main ethnic group. Results showed that more than three-quarters of the multi-ethnic sample in each age group selected a single main ethnic group when required. The proportion that selected one ethnic group was higher in the adolescent sample (96%), who had one alternative option of "I can't choose only one ethnic group", compared to children (by mother proxy) and adults (77% and 80%, respectively), who were given the option to select up to two main ethnic groups, "don't know", or refuse response. Interestingly, the proportion of children and adults who selected one main ethnic group was more comparable to previous research with early adolescents in Aotearoa New Zealand that provided five alternative response options including "depends on who with", "no main ethnic group", and "don't know main ethnic group" (74%; Kukutai & Callister, 2009). It is therefore possible that more numerous and open response options may result in a lower rate of self-prioritisation compared to more restrictive options, but more systematic research is needed to test this hypothesis.

The ability of the majority of multi-ethnic participants to self-select a main ethnic group is supported by findings in the United States with nationally representative adolescent and adult data (D. R. Harris & Sim, 2002; Ingram et al., 2003). In general, research to date suggests that self-prioritisation can be a valid method of classifying multiple ethnic identifications into a single category if it is required and appropriate to the research question. However, it is also important to consider those who refused to select a main ethnic group. Aspinall and Song's (2013) qualitative investigation of multi-ethnic university students in Britain found that

participants who refused response expressed that they were uncomfortable with denying a part of their heritage, or rejected the notion of traditional ethnic categories altogether. This demonstrates that asking for self-prioritised ethnicity can be problematic from an ethical perspective. It is a particularly pertinent concern in Aotearoa New Zealand, because the concept of not acknowledging all whakapapa (ancestry) is considered offensive in Māori culture (Jackson, 2003; Khawaja et al., 2000).

The second research question examined the rate of alignment between self-prioritised and administratively-prioritised ethnicity for multi-ethnic participants in each age group. Consistent with previous studies (e.g., Atatoa Carr et al., 2017; Kukutai & Callister, 2009), the results showed a relatively high level of discrepancy between the two prioritisation methods. In this study, the discrepancy rate between the two prioritisation methods was around 60% in adults. This is higher than the rate previously reported in Atatoa Carr et al.'s (2017) analysis of the same adult Growing Up in New Zealand dataset (approximately 40%), because Atatoa Carr et al.'s calculations excluded participants who did not select a main ethnic group. The current study classified these cases as being discrepant with administratively-prioritised ethnicity because it yields a different result than the participant's response. The current study also extended the analyses to nationally representative samples of children and adolescents. The discrepancy rate in the adolescent sample was also around 60%; however, a higher discrepancy rate of almost 70% was observed in the child sample.

The overall high discrepancy rate between self-prioritisation and administrative-prioritisation of multiple ethnic identifications found across all age groups contributes empirical evidence to the largely theoretical argument that administratively-prioritised ethnicity is a poor measure of a multi-ethnic individual's strength of ethnic self-affiliation (Statistics New Zealand, 2004). However, this finding must be interpreted with caution because, for ethnic minorities, racism and social stigma may influence self-prioritised ethnicity (Herman,

2004). Although the particular prioritisation hierarchy examined in this study is specific to Aotearoa New Zealand, it is likely that external prioritisation algorithms used in other countries (e.g., the deterministic whole allocation methods used in the United States based on the relative size of racial groups; OMB, 2000) will also produce a high discrepancy rate with self-prioritised ethnicity. The discrepancy between the two prioritisation methods and the differing rationale behind each method indicates that the validity of administrative-prioritisation and self-prioritisation needs to be assessed in light of each research or policy issue under investigation.

### 4.5.2 Consistency of Ethnic Classification Methods

The third research question examined the consistency of ethnic group proportions in each age group across different ethnic classification methods at the overall sample level (i.e., including mono-ethnic participants). Three relatively accessible classification methods which can output multiple ethnicities into a small number of broad ethnic groupings were compared: total response, administrative-prioritisation, and self-prioritisation. As expected, both methods of prioritisation understated ethnic group proportions compared to total response, except for administratively-prioritised Māori, because Māori is in first position in the prioritisation hierarchy as per Indigenous Treaty rights (Department of Statistics, 1993). Similar to previous analyses that compared total response to administrative-prioritisation using census data from Aotearoa New Zealand (Callister et al., 2007), the current study found a larger percentage of understatement by both administrative-prioritisation and self-prioritisation in subgroups with higher rates of multi-ethnicity (e.g., children and adolescents). In this study, the magnitude of understatement tended to be twice as large in children and adolescents than in adults. Total response ethnicity is therefore problematic from an equity perspective because it minimises the influence of Indigenous and ethnic minority groups.

The current research also investigated the differential effect of administrative-prioritisation and self-prioritisation on ethnic group proportions in each age group at the overall sample level. These two methods have previously only been compared in samples consisting solely of multi-ethnic participants (e.g., Atatoa Carr et al., 2017; Kukutai & Callister, 2009). Results showed substantial discrepancy between the proportions outputted by administrative-prioritisation and self-prioritisation for both Europeans and Māori. In each age group, compared to the respective proportion outputted by self-prioritisation, the proportion outputted by administrative-prioritisation was smaller in European and larger in Māori. The magnitude of effect was again around twice as large in young people, who are less likely to be influenced by socially ascribed notions of ethnicity (Perez & Hirschman, 2009). This supports the finding from the OMB's (2000) simulation study that racial classification methods in the United States have a stronger effect when there are higher levels of racial diversity. Overall, inclusion of the more ethnically diverse samples of children and adolescents indicates that researchers' choice of ethnic classification method will have an increasing impact on population-level ethnic distributions over time, and that this impact will be particularly pronounced for ethnic groups with higher rates of multi-ethnicity, such as Māori.

### 4.5.3 Limitations

Although the use of large-scale datasets of children, adolescents, and adults that included information on both total response and self-prioritised ethnicity is a strength of this study, some limitations should be noted. First, while the child sample was broadly generalisable to live births in Aotearoa New Zealand (Morton et al., 2015), and the adolescent sample was a nationally representative sample of secondary school students (Clark et al., 2013), the adult sample used in this study was recruited during the mothers' pregnancy with the child and therefore was not representative of the wider adult population in Aotearoa New Zealand (Morton et al., 2013). However, the adult dataset is still valuable because of its large sample

size and self-prioritised ethnicity data. Second, differences in survey methodology between the age groups mean that age comparisons need to be interpreted with care. For example, adolescent and adult responses were self-reported, but child data were collected by mother proxy. This is a common method of collecting children's ethnicity data, but it may not necessarily reflect children's self-identification (Perez, 2006; Statistics New Zealand, 2005). Third, ethnicity survey questions and response options for adolescents were different than those for children and adults (see Table 4.2). The aggregation of responses into broad ethnic groupings mitigated some comparability issues.

It should also be noted that the focus of this study was cross-sectional measurement of ethnicity aggregated at a broad level of ethnic grouping, as this is the most common way of collecting and outputting ethnicity data (OMB, 2000; Statistics New Zealand, 2005). Therefore, heterogeneity in self-prioritisation within broad ethnic groupings was noted but not explored in depth. Finally, previous research indicates that ethnic identification can be fluid across time and contexts (D. R. Harris & Sim, 2002; Liebler et al., 2017), so ethnicity responses cannot be interpreted as a static characteristic of participants.

### 4.5.4 Implications

Notwithstanding these limitations, the current study strongly demonstrates that researchers' choice of ethnic classification method can have a significant influence on ethnic group proportions and therefore ethnic group counts, especially in populations with higher rates of multi-ethnicity, such as children and adolescents. Accurate ethnic counts are crucial to research and policy aimed at monitoring responsibilities to Indigenous Peoples and reducing ethnic inequities in health and social outcomes (Balestra & Fleischer, 2018; Mays et al., 2003). Therefore, with the increasing prevalence of multiple ethnic identifications due to interethnic unions and changing social understandings of identity (Aspinall, 2018b; Balestra & Fleischer, 2018; Perez & Hirschman, 2009; Statistics New Zealand, 2014b), ethnic classification method

is a pertinent issue in Aotearoa New Zealand and other multi-ethnic contexts. This issue poses the complex question of which ethnic classification method is the best to use. The current research illustrates the importance of choosing the most appropriate method for the specific research question or policy problem.

Three relatively accessible ethnic classification methods that can output multiple ethnic identifications into broad ethnic groupings were compared in this study: total response, administrative-prioritisation, and self-prioritisation. Statistics New Zealand (2004) recommends the use of total response ethnicity, and sole/combination grouping when more nuance is needed. The strength of the all-inclusive total response method is that it retains all the ethnic groups that an individual identifies with. However, it increases the influence of the dominant ethnic group and minimises the influence of Indigenous and other minority ethnic groups, which may be problematic when used for purposes such as resource allocation. Furthermore, because the outputted groups are not mutually exclusive, it is not suitable for some statistical techniques, and results can be difficult to interpret for both researchers and the research audience alike. This classification method can also substantially over-count multi-ethnic individuals. For example, the number of total ethnicity responses of children and adolescents in this study exceeded their respective sample sizes by up to 40%.

For these reasons, researchers generally prefer to work with mutually exclusive ethnic categories (Callister et al., 2007; Mays et al., 2003). Administrative-prioritisation provides a simple way to externally classify multiple responses into single categories, and is widely used in health and education research in Aotearoa New Zealand (Education Counts, 2014; Ministry of Health, 2017). Because it prioritises the Māori ethnic group, it is suited to resource allocation and policy development in the Aotearoa New Zealand context as part of Indigenous Treaty obligations. However, the discrepancy between administratively-prioritised and self-prioritised ethnicity found in this study suggests that there are limitations to the administrative method if

it is used as a proxy for multi-ethnic participants' strongest ethnic self-affiliation. This study indicates that, in general, asking participants to self-select their main ethnic group appears to be a valid method to prioritise multiple ethnicities if strength of self-affiliation is important. However, shortcomings of the self-prioritisation approach should be noted. In particular, responses may fluctuate across social contexts or be influenced by developmental stage, racism, and social stigma (D. R. Harris & Sim, 2002; Herman, 2004); and asking for a self-prioritised ethnicity may be ethically contentious or offensive to some participants (Aspinall & Song, 2013; Sanchez, 2010; Townsend et al., 2009).

Given that each ethnic classification method has strengths and limitations, it is apparent that the "best" classification method depends on the purpose of each individual study or policy problem. For example, if it is important in a study that all ethnic affiliations of a multi-ethnic participant are captured, then total response may be more suitable. If additional nuance is needed, sole/combination grouping may be endorsed. If strength of self-identification and mutually exclusivity are required, self-prioritised ethnicity may be more appropriate. If affiliation with certain ethnic groups is relevant (e.g., for policy or funding purposes), administrative-prioritisation may be considered. It is paramount that researchers collect appropriate and quality ethnicity data that allow the most suitable classification method to be used (Aspinall, 2018b).

## 4.6 Conclusion

The significant effect of ethnic classification method on ethnic group counts, especially when there is a sizeable multi-ethnic proportion, necessitates that researchers working with ethnicity data in multi-ethnic contexts critically engage in the decision process regarding the classification of multiple ethnic responses. The most appropriate method for the research question should be selected and clearly documented in research dissemination alongside the

rationale behind the decision. Possible implications of the choice of ethnic classification method also need to be explicitly and transparently considered and discussed.

CHAPTER 5

# Study 2: Context Matters

Study 1 in the previous chapter showed that ethnic classification method matters with respect to overall ethnic group size, particularly in samples with higher multi-ethnic identification rates (e.g., children and adolescents). Hence, in the global context of steadily increasing multi-ethnic identification (Aspinall, 2018b; N. Jones et al., 2021; Statistics New Zealand, 2020a), researchers are in a position of power, because their ethnic classification choice affects who is included or excluded from outputted ethnic groups. Another important finding from Study 1 is the high rate of discrepancy between administrative-prioritisation and self-prioritisation in each age cohort (≥60%). These two prioritisation methods are popular among applied researchers because they output multiple ethnic responses into a small number of mutually exclusive ethnic groups—a characteristic that is particularly attractive for inferential statistical techniques (e.g., ANOVA, regression, and SEM; Cohen et al., 2003; Field et al., 2012). However, there is scant information on whether these discrepancies are randomly distributed, or systematically associated with demographic characteristics (e.g., age, sex, birthplace, and socioeconomic deprivation). If discrepancies between administrative-prioritisation and self-prioritisation are systematically associated with demographic characteristics, then researchers' choice of prioritisation method can impact the demographic composition of their sample, and potentially bias results and conclusions.

Therefore, Study 2 (this chapter) investigates how individual characteristics (e.g., ethnic combination, sex, and birthplace) and contextual characteristics (e.g., neighbourhood socioeconomic deprivation and ethnic composition) are associated with discrepancies between administratively-prioritised and self-prioritised ethnicity in multi-ethnic children, adolescents,

and adults. Two types of discrepancies will be explored: (1) discrepancies due to participants selecting a main ethnic group that differs from their administratively-prioritised ethnicity, and (2) discrepancies due to participants opting not to select a main ethnic group. This will inform researchers of the effects that their choice of prioritisation method can have on the demographic composition of their sample.

The study is published in *Social Science Research*, and is presented in its entirety below, with minor edits for consistency. The recommended citation for the article is:

Yao, E. S., Meissel, K., Bullen, P., Clark, T. C., Atatoa Carr, P., Tiatia-Seath, J., Peiris-John, R., & Morton, S. M. B. (2022). Demographic discrepancies between administrative-prioritisation and self-prioritisation of multiple ethnic identifications. *Social Science Research*, *103*, 1–16. https://doi.org/10.1016/j.ssresearch.2021.102648

## 5.1 Introduction

Ethnic classification for research purposes may initially appear objective and neutral, but is inherently a subjective and value-laden process that can include or exclude individuals from group membership (Kertzer & Arel, 2002; Smedley & Smedley, 2005). The steadily increasing global multi-ethnic population, driven by transnational mobility, interethnic unions, and changing patterns of self-identification, adds complexity to ethnic classification because traditional ethnic group boundaries are increasingly blurred (Callister et al., 2007; Khanna, 2012; Morning, 2008). There are several classification options available to researchers, each with differing ease of statistical application and ethical implications (see Yao et al., 2021). Because different ethnic classification methods can significantly affect the relative size of ethnic groups and the interpretation of ethnic differences in health and social outcomes (Callister et al., 2007; Mays et al., 2003; Yao et al., 2021), researchers are in a position of power when making ethnic classification decisions.

There is no universal "best" ethnic classification method, as this depends on the research question and sociocultural context (Atatoa Carr et al., 2017; Cormack & Robson, 2010; Yao et al., 2021). However, for multivariable statistical analysis, applied researchers typically select classification methods that output multiple ethnic identifications into a small number of mutually exclusive categories (Callister et al., 2007; Mays et al., 2003; Yao et al., 2021). Two common methods are administrative-prioritisation (a coding scheme which prioritises multiple ethnic responses into a single ethnic group) and self-prioritisation (obtained via a follow-up survey question which asks multi-ethnic participants to self-select a "main" ethnic group). Previous research indicates considerable discrepancy in how multi-ethnic participants are classified according to these two methods (Atatoa Carr et al., 2017; Kukutai & Callister, 2009; Yao et al., 2021), but scant research has investigated whether there are systematic demographic differences underlying these discrepancies, and whether these differ across developmental stages (e.g., childhood, adolescence, and adulthood). If systematic differences exist, then researchers' choice of prioritisation method may affect the demographic composition of outputted ethnic groups and potentially influence or bias the results. This has important implications for knowledge construction and policy decisions. For example, if one prioritisation method is more likely to classify multi-ethnic participants living in higher socioeconomic deprivation to an ethnic minority group, the conflation of ethnicity and deprivation can result in relatively larger associations between ethnicity and health outcomes, and relatively smaller associations between deprivation and health outcomes. This can potentially lead researchers and policymakers to pay less attention to the detrimental impact of deprivation than what is warranted.

To address this gap, the current study will utilise large-scale survey data of children, adolescents, and adults from Aotearoa New Zealand to empirically investigate whether discrepancies between administrative-prioritisation and self-prioritisation of multiple ethnic

identifications are associated with key demographic characteristics such as age, sex, education level, socioeconomic deprivation, and neighbourhood ethnic composition. We acknowledge that prioritisation methods are inherently limited because they suppress multi-ethnic identifications. However, given the ubiquity, popularity, and accessibility of prioritisation, we consider this an important starting point for extending critical discussion on researchers' choice of ethnic classification method, both in Aotearoa New Zealand and in other multi-ethnic English-speaking countries (e.g., Australia, Canada, the United States, and United Kingdom).

### 5.1.1 Conceptual Background

Ethnicity is a social construct used to characterise a group of people who share commonalities in ancestry, history, culture, traditions, values, and beliefs (Smedley & Smedley, 2005). It is considered a multidimensional construct comprising three overlapping dimensions: ancestry, self-identification, and observer-identification (Roth, 2016; Woo et al., 2011). Ethnicity originated from the biological concept of race, which gained popularity in the 18$^{th}$ century due to attempts to justify colonisation and slavery (Smedley & Smedley, 2005). Supported by racial (pseudo-)science at the time, the traditional concept of race organised humans hierarchically based on innate physical characteristics stemming supposedly from genetic composition (e.g., skin colour). Historically, race was typically classified in a top-down manner by colonised states to control and dominate (e.g., in the United States, the "one-drop rule" limited multiracial African Americans to "hypodescent" to prevent assimilation and maximise the number of slaves [Smedley & Smedley, 2005]; while in Aotearoa New Zealand, earlier classification as Māori, the Indigenous Peoples of the country, was restricted to those who reported at least "half" Māori descent to expedite assimilation and land appropriation [Cormack & Robson, 2010; Rocha, 2012]). However, subsequent developments in genetic research disproved the biological premise for race as limited genetic between-group differences were found (Smedley & Smedley, 2005). Following World War II, race began evolving into the

social construct now referred to as "ethnicity" in most English-speaking countries outside of the United States,[22] and from the 2000 global census round, is typically collected via self-identification using a check-all-that-apply approach (Morning, 2008).

Ethnicity is of present-day importance because it is essential to ethnic group recognition as well as equity and antidiscrimination agendas (Atatoa Carr et al., 2017; Mays et al., 2003; Smedley & Smedley, 2005; Statistics New Zealand, 2005). Contemporary researchers and policymakers regularly use ethnicity to describe and monitor health and social inequities arising from longstanding structural racism experienced by Indigenous and ethnic minority groups. Ethnicity is also commonly used as a proxy for unmeasured sociocultural and contextual processes underlying these inequities, although ideally the process(es) of interest (e.g., socioeconomic deprivation, perceived racism, and experiences of discrimination) should be measured directly (Mays et al., 2003; Moubarac, 2013). Irrespective of how it is used, ethnic classification is clearly a prerequisite for any research involving ethnicity.

### 5.1.2 Contextual Background

Aotearoa New Zealand is a valuable context for research on ethnic classification because of its long history of interethnic marriage, high levels of ethnic diversity, and longstanding official collection of self-identified multi-ethnic responses (Callister et al., 2007; Cormack & Robson, 2010; Khawaja et al., 2000). Further, it is obliged under its founding document, *Te Tiriti o Waitangi* (*The Treaty of Waitangi*),[23] to monitor and address ethnic inequities. Unlike many other colonised states, intermarriage between Māori and Europeans was never legally prohibited, but instead was considered a method to "civilise", assimilate, and

---

[22] "Ethnicity" (rather than "race") is used throughout this paper except when it pertains to the U.S. context.

[23] *Te Tiriti o Waitangi* was signed in 1840 and guarantees equal rights between Māori and non-Māori.

acquire land from Māori (Cormack & Robson, 2010; Rocha, 2012). This agenda was reflected in census collections in the first half of the 1900s, when respondents were asked to indicate their descent by "blood quantum" (e.g., "halves", "quarters", "eighths"). In contrast to traditional Māori culture which defines group belonging based on whakapapa (ancestry) regardless of "blood quantum" (Jackson, 2003; Khawaja et al., 2000), prior to 1974 the colonial government only classified those who reported at least "half" Māori descent as Māori (Cormack & Robson, 2010; Khawaja et al., 2000).

In the second half of the 1900s, Māori rights movements and changing understandings of ethnicity led to a substantial modification of the ethnicity question in the five-yearly census. The revised question, first introduced in 1986, asked respondents to self-identify their ethnic origin(s) using a check-all-that-apply approach ("ethnic origin(s)" was replaced with "ethnic group(s)" in 1991; Cormack & Robson, 2010; Khawaja et al., 2000). Around this time, there was a wave of migration from neighbouring Pacific nations (particularly Sāmoa, Tonga, and Fiji), followed by migrants from Asia (although a small number of Chinese had settled earlier in the 1860s Gold Rush), and later, migrants and refugees from Middle Eastern, Latin American, and African (MELAA) countries. As of the 2018 Census, the main ethnic groupings in Aotearoa New Zealand in descending total response size were: European (70%), Māori (17%), Asian (15%), Pacific Peoples (8%), and MELAA (2%; total percentage exceeds 100% because respondents could select multiple ethnicities; Statistics New Zealand, 2020a). Note these are broad ethnic groupings utilised for statistical output, and there is considerable heterogeneity within each group. Overall, 11% reported more than one broad ethnic grouping, with this being more common in Māori, Pacific, and younger generations.

Statistics New Zealand's (2005) *Statistical Standard for Ethnicity* provides guidelines on the official collection and classification of ethnicity data. It prescribes two standard classification options for presenting ethnicity data: (1) *total response ethnicity*, which counts

each respondent in all of their reported ethnic groups (e.g., "European alone or in combination", "Māori alone or in combination", etc.); and (2) *sole/combination ethnicity*, which counts each respondent once according to their reported single ethnic group (e.g., "European only", "Māori only", etc.) or combination of groups (e.g., "Māori/European", "Pacific/European", etc.). However, these methods can be difficult to use in multivariable statistical analysis because they do not output multiple ethnic responses into a small number of mutually exclusive categories (Callister et al., 2007; Mays et al., 2003; Yao et al., 2021). Therefore, ethnic prioritisation methods tend to be more popular among applied researchers (Moubarac, 2013; Yao et al., 2021).

### 5.1.3 Ethnic Prioritisation Methods

*Administrative-prioritisation* (prioritisation by the researcher) and *self-prioritisation* (prioritisation by the participant) are two common prioritisation methods used to classify multiple ethnic identifications into a single ethnic group. In the administrative-prioritisation approach, multi-ethnic participants are allocated to the ethnic group in the highest position of a predetermined hierarchy. In Aotearoa New Zealand, the following hierarchy was introduced as the official ethnic classification method by the Department of Statistics (1993; now Statistics New Zealand): Māori > Pacific Peoples > Asian > other ethnic groups except for European > European. This prioritises Māori in recognition of Treaty rights, followed by numerically smaller ethnic groups, so that they are not subsumed within the dominant New Zealand European majority. In the United States, administrative-prioritisation (termed "deterministic whole allocation") is usually based on the relative size of racial groups according to census data (allocation schemes include "Smallest Group", "Largest Group", and "Largest Group Other Than White"; Mays et al., 2003).

Compared to self-prioritisation, the key advantages of administrative-prioritisation are its ability to: (1) prioritise multiple ethnic responses without seeking additional information; (2) elevate Indigenous and marginalised groups for equity, antidiscrimination, and affirmatory

action reasons; and (3) ensure that smaller subgroups are not subsumed by the dominant group (Department of Statistics, 1993; Yao et al., 2021). Its main limitation is its dismissal of the self-affiliation of multi-ethnic participants, and for this reason, has been discontinued as the official classification method by Statistics New Zealand (2005). However, ostensibly due to the aforementioned advantages, it remains the standard ethnic classification method in sectors where equity and resourcing are important concerns (e.g., education and health; Education Counts, 2014; Ministry of Health, 2017), and continues to be widely used by researchers (Moubarac, 2013).

In the self-prioritisation approach, multi-ethnic participants are asked in a follow-up survey question to self-select the ethnic group that they identify with most (commonly called "main" ethnic group in Aotearoa New Zealand, or "best" single race in the United States; D. R. Harris & Sim, 2002; Kukutai & Callister, 2009). This method, usually observed in sample surveys, appears to shift the locus of choice to participants and can arguably be a useful indicator of strength of ethnic self-affiliation. Previous research in Aotearoa New Zealand and overseas indicates that the majority of multi-ethnic participants will select a main ethnic group when required (Aspinall & Song, 2013; Campbell, 2007; D. R. Harris & Sim, 2002; Kukutai & Callister, 2009; Yao et al., 2021). However, self-prioritisation has also been termed the "forced choice dilemma" (Sanchez, 2010; Townsend et al., 2009) because it can be a complex and cognitively demanding question that forces individuals to prioritise one of their ethnic groups over another, and in the process deny a part of their identity. It can convey to participants that multi-ethnic identity is undervalued, and has been found in experimental studies to be associated with negative outcomes such as higher depressive symptoms and lower self-esteem (Sanchez, 2010; Townsend et al., 2009). The ethical concerns associated with self-prioritisation is particularly pertinent in the Aotearoa New Zealand context because, as previously explained,

whakapapa (ancestry) is considered indivisible in Māori culture, so participants may find the self-prioritisation question offensive (Jackson, 2003; Khawaja et al., 2000).

Further, although participants may appear to be "free" to select the ethnic group that best represents them, in reality self-prioritisation is influenced by an interaction of personal, social, and structural influences in conjunction with survey methodology (see Khanna, 2012). On a personal level, participants' ancestry, birthplace, nationality, age, physical appearance, language proficiency, personal conceptualisation of ethnicity, and perception of how others see them, may all influence their response (Callister et al., 2007; Herman, 2004; Khanna, 2012; Schenker & Parker, 2003). On a social level, family ethnic socialisation, peer groups, and neighbourhood ethnic composition have been shown to be associated with responses via cultural exposure and experiences (Campbell, 2007; Herman, 2004; Khanna, 2012). On a structural level, responses may be influenced by sociohistorical background, stigmatisation, and racism (Campbell, 2007; Herman, 2004; Roth, 2005). In addition, survey format (e.g., wording and response options), context (e.g., home vs. school), and participants' beliefs about why ethnic data are being collected (e.g., to access opportunities vs. to marginalise) may all affect self-prioritisation (Callister et al., 2007; D. R. Harris & Sim, 2002).

### 5.1.4 Discrepancies Between Prioritisation Methods

Previous studies in Aotearoa New Zealand have found over a 40% discrepancy rate between administratively-prioritised and self-prioritised ethnicity (Atatoa Carr et al., 2017; Kukutai & Callister, 2009; Yao et al., 2021), indicating that the two methods cannot be regarded as interchangeable. From a theoretical perspective, researchers argue that different classification methods may be suitable for different research purposes (Roth, 2016; Woo et al., 2011; Yao et al., 2021). For example, administrative-prioritisation may be better suited for research where there are specific ethnic groups of interest, whereas self-prioritisation may be more suitable when strength of ethnic self-affiliation is important. However, despite the

ubiquitous use of ethnic prioritisation methods in Aotearoa New Zealand and other English-speaking countries, there is a lack of comprehensive knowledge around how researchers' choice of prioritisation method affects the demographic composition of outputted ethnic groups, and how differing group composition potentially impacts research and policy aiming to address ethnic inequities.

While we could not identify any empirical studies which explicitly examined demographic characteristics associated with discrepancies between ethnic prioritisation methods, we identified one study in Aotearoa New Zealand, and a small body of research in the United States, which examined demographic characteristics associated with multi-ethnic participants' self-prioritisation of their main ethnic group (e.g., age, educational level, socioeconomic status, neighbourhood ethnic composition). The Aotearoa New Zealand study by Kukutai (2004) focused on women who identified as both Māori and European in the 1995 New Zealand Women: Family, Employment and Education survey. Bivariate analysis indicated that those who self-prioritised as European ($n$ = 77) were more likely than those who self-prioritised as Māori ($n$ = 67) to have higher educational attainment, higher personal income, and live in areas of lower Māori composition (Kukutai, 2004). No age differences were observed. The study excluded those who did not select a main ethnic group ($n$ = 39; note this indicates that 21% of the sample were not able or willing to self-prioritise a main ethnic group).

In the United States, large-scale datasets (e.g., the National Health Interview Survey and the National Longitudinal Study of Adolescent to Adult Health) have been utilised to explore demographic characteristics associated with participants' choice of main race. Results for common racial combinations in adults (e.g., Campbell, 2007; Schenker & Parker, 2003), adolescents (e.g., D. R. Harris & Sim, 2002; Herman, 2004), and parents' identification of children (e.g., Bratter, 2007; Brunsma, 2005; Roth, 2005), indicate that participants' main race tends to align with the race most prevalent in their neighbourhood. In addition, higher

socioeconomic status, typically measured through family income or a proxy measure such as highest education, was associated with a higher likelihood of selecting a "higher status" main race (e.g., non-Hispanic White) or a multiracial identity.

Overall, it can be inferred from existing studies that neighbourhood ethnic composition and socioeconomic status will likely be associated with discrepancies between administratively-prioritised and self-prioritised ethnicity. However, more comprehensive research is needed to investigate how demographic characteristics are systematically related to discrepancies that arise due to the prioritisation method chosen, and whether this differs across age cohorts. This will inform researchers of the potential implications of their decisions, and is particularly important given the steadily increasing rate of multi-ethnic identification.

## 5.2 The Current Study

Therefore, in the current study, we investigate the following research question in three age cohorts (children, adolescents, and adults) in Aotearoa New Zealand: *what demographic characteristics are associated with discrepancies between administratively-prioritised and self-prioritised ethnicity for those who select more than one ethnic group?* Two types of discrepancies will be examined: (1) discrepancies arising from participants' selection of a main ethnic group that was *different* from administrative-prioritisation, and (2) discrepancies due to participants' *non-selection* of a main ethnic group. Due to the exploratory nature of this study, we will focus on investigating demographic characteristics that are commonly used in multivariable analyses (e.g., age, sex, education level, socioeconomic deprivation, and neighbourhood ethnic composition).

## 5.3 Method

### 5.3.1 Data Sources

Secondary data from the Growing Up in New Zealand study was utilised for the child and adult samples, and the Youth'12 National Youth Health and Wellbeing Survey for the adolescent sample. Growing Up in New Zealand is a longitudinal birth cohort study which follows a broadly nationally representative sample of over 6,000 children and their parents from the mother's pregnancy (see Morton et al., 2013). Pregnant mothers were recruited between 2009 and 2010 in the ethnically diverse regions of Auckland and Waikato in the North Island of Aotearoa New Zealand. Mother data from the antenatal data wave ($N$ = 6,822) were utilised for the adult sample, and 54-month child data ($N$ = 6,156) were utilised for the child sample. These data waves respectively contained the most detailed ethnicity data for these age groups. The antenatal data wave also surveyed mothers' partners (typically children's biological fathers), but partner data were not utilised due to missing meshblock information (see Section 5.3.2). Adult data were collected using face-to-face computer-assisted personal interviews, and child data were collected by proxy from the child's mother (or main caregiver) using the same method. Ethical approval for Growing Up in New Zealand was obtained from the Ministry of Health Northern Y Regional Ethics Committee (reference NTY/08/06/055).

The Youth'12 dataset ($N$ = 8,500), used for the adolescent sample, was a nationally representative cross-sectional survey of secondary school students (aged 12–18 years) in 91 randomly selected schools across Aotearoa New Zealand (see Clark et al., 2013). Students in smaller schools were oversampled to protect confidentiality, and sampling weights were used to adjust for selection likelihood. The survey was administered in schools in 2012 via an online questionnaire on computer tablets. Ethical approval for Youth'12 was obtained from the University of Auckland Human Participants Ethics Committee (reference 2011/206).

The analytic samples for this study were selected from the child, adolescent, and adult datasets in three successive steps. First, participants were excluded if they identified with only one broad ethnic grouping (i.e., European, Māori, Pacific, Asian, MELAA, or Other; excluded child $n$ = 3,781, adolescent $n$ = 5,777, adult $n$ = 5,723). Second, participants who identified with the European/Other combination were excluded (excluded child $n$ = 305, adolescent $n$ = 242, adult $n$ = 40). This was because open-ended responses for "Other" revealed that the majority identified as "New Zealander", and New Zealander is arguably a nationality rather than an ethnicity. Third, participants with missing demographic data (see Section 5.3.2) were excluded (excluded child $n$ = 210, adolescent $n$ = 68, adult $n$ < 10). The proportion of missingness (≤10%) was within the range generally considered as unproblematic if omitted (Schlomer et al., 2010). The final analytic sample sizes for children, adolescents, and adults were 1,860; 2,413; and 1,056; respectively.

Demographic information for each sample is presented in Table 5.1. Demographic information by ethnic combination is available in Supplementary Tables B.1 to B.3 in Appendix B. Note the child and adult samples were not entirely independent, as around 70% of the final adult sample ($n$ = 763) were mothers of children in the final child sample. Missing child observation was the most common reason why mothers did not have children in the child sample.

### 5.3.2 Measures

Demographic information available for each dataset included total response ethnicity, self-prioritised ethnicity, sex, age, birthplace, and meshblock code (a small geographic unit derived from participants' address). Total response ethnicity was obtained by asking participants to select *all* the ethnic groups they identified with, and self-prioritised ethnicity was obtained in a follow-up question which asked participants who selected more than one

**Table 5.1**

*Demographic Characteristics by Sample*

| | Children[a] (*N* = 1,860) | | Adolescents (*N* = 2,413) | | Adults (*N* = 1,056) | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| Ethnic combination[b] | | | | | | |
| Māori/European | 797 | 43 | 1,071 | 44 | 627 | 59 |
| Pacific/European | 202 | 11 | 418 | 17 | 130 | 12 |
| Asian/European | 162 | 9 | 237 | 10 | 44 | 4 |
| MELAA/European | 63 | 3 | 88 | 4 | 25 | 2 |
| Māori/Pacific | 133 | 7 | 78 | 3 | 72 | 7 |
| Pacific/Asian | 26 | 1 | 103 | 4 | 15 | 1 |
| Asian/Other | 82 | 4 | 38 | 2 | 18 | 2 |
| Māori/Pacific/European | 190 | 10 | 138 | 6 | 63 | 6 |
| Other combinations | 205 | 11 | 242 | 10 | 62 | 6 |
| Sex | | | | | | |
| Female | 897 | 48 | 1,295 | 54 | 1,056 | 100 |
| Male | 963 | 52 | 1,118 | 46 | - | - |
| Age group (years) | | | | | | |
| Children | | | | | | |
| 4–5 | 1,860 | 100 | - | - | - | - |
| Adolescents | | | | | | |
| <16 | - | - | 1,654 | 69 | - | - |
| ≥16 | - | - | 759 | 31 | - | - |
| Adults | | | | | | |
| <30 | - | - | - | - | 624 | 59 |
| ≥30 | - | - | - | - | 432 | 41 |
| Birthplace | | | | | | |
| Aotearoa New Zealand | 1,860 | 100 | 1,968 | 82 | 935 | 89 |
| Overseas | - | - | 445 | 18 | 121 | 11 |
| Highest educational attainment[c] | | | | | | |
| Secondary school qualification or below | 673 | 36 | - | - | 432 | 41 |
| Certificate or diploma | 613 | 33 | - | - | 365 | 35 |
| Bachelor's or above | 574 | 31 | - | - | 259 | 25 |
| Urbanicity | | | | | | |
| Main urban | 1,520 | 82 | 1,866 | 77 | 862 | 82 |
| Non-main urban | 199 | 11 | 262 | 11 | 122 | 12 |
| Rural | 141 | 8 | 284 | 12 | 72 | 7 |
| Socioeconomic deprivation (NZDep) | | | | | | |
| Low dep. (1–3) | 419 | 23 | 438 | 18 | 186 | 18 |
| Medium dep. (4–7) | 659 | 35 | 858 | 36 | 352 | 33 |
| High dep. (8–10) | 782 | 42 | 1,117 | 46 | 518 | 49 |
| Area ethnic similarity (%)[d] | | | | | | |
| Mean (*SD*) | 18.15 (13.99) | | 22.09 (18.46) | | 19.27 (13.73) | |
| Skewness | 1.40 | | 1.11 | | 1.36 | |
| Kurtosis | 2.11 | | 0.47 | | 2.24 | |

*Note.* A hyphen indicates that the variable was not applicable for the sample. MELAA = Middle Eastern, Latin American, and African.

[a]Child responses were collected by proxy from the child's mother.

[b]Combinations with $n < 50$ were aggregated into "other combinations" for analysis.

[c]Mother's highest educational attainment used for children.

[d]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; unstandardised.

ethnic group to indicate their *main* ethnic group. Both variables were collected via a detailed list of ethnic categories (e.g., New Zealand European, English, Australian, Māori, Sāmoan, Tongan, Chinese, Indian, Middle Eastern, Latin American, Other). Responses were aggregated into six broad ethnic groupings (i.e., European, Māori, Pacific, Asian, MELAA, Other). For the self-prioritisation question, the Youth'12 survey provided an additional "I can't choose only one ethnic group" option, whereas the Growing Up in New Zealand survey allowed participants to select up to two main ethnic groups, "don't know", or refuse response. Administratively-prioritised ethnicity was coded from total response ethnicity using Department of Statistics' (1993) hierarchy: Māori > Pacific > Asian > MELAA > Other > European.

#### *5.3.2.1 Dependent Variable*

The dependent variable had three categories indicating the type of discrepancy between administratively-prioritised and self-prioritised ethnicity when aggregated into six broad ethnic groupings:

- 0 = *non-discrepant* (participants' self-prioritised ethnicity was *identical* to their administratively-prioritised ethnicity; treated as the reference group),
- 1 = *discrepant—different main ethnic group* (participants self-prioritised a main ethnic group that *differed* from their administratively-prioritised ethnicity), and,
- 2 = *discrepant—no main ethnic group* (participants did *not* self-prioritise one main ethnic group [e.g., selected two, could not choose, or refused response], so it was discrepant from their administratively-prioritised ethnicity).

#### *5.3.2.2 Individual-Level Independent Variables*

All individual demographic characteristics were examined as dummy-coded categorical variables. *Ethnic combination* was participants' specific combination of broad total response ethnicities. Combinations comprising both European and Other (mainly New Zealander

responses) were simplified to European only, as both reflected identification with the dominant group (e.g., Māori/European/Other was classified as Māori/European). However, Other was not recoded to European when it was not paired with it (e.g., Asian/European and Asian/Other were treated as distinct combinations), because European implies European ancestry whereas Other may not (e.g., second-generation Asian-New Zealanders). Within each sample, ethnic combinations with at least 50 participants were specified as independent categories; remaining combinations were aggregated as "other combinations". Māori/European was the largest combination in each sample (>40%) and was treated as the reference category.

*Sex*, *age group*, *birthplace*, and *highest educational attainment* were also examined where applicable to the sample. For example, age group and birthplace were not examined among children as they were approximately the same age and all born in Aotearoa New Zealand, whereas sex was not examined among adults as all were female. Highest educational attainment was not relevant to children and adolescents as they had not completed compulsory schooling. Instead, maternal highest education was used for children. Maternal education level was not available for adolescents.

#### *5.3.2.3 Contextual-Level Independent Variables*

All contextual demographic characteristics were based on meshblock code derived from participants' home address. Meshblock is the smallest geographic unit defined by Statistics New Zealand (2017), ranging from approximately 30 to 60 dwellings (60–120 residents). *Urbanicity* was coded into three levels: main urban (population ≥ 30,000), non-main urban (population between 1,000 and 29,999), and rural (population < 1,000; Statistics New Zealand, 2017). *Area socioeconomic deprivation* was measured by the New Zealand Deprivation Index (NZDep; Atkinson et al., 2014). NZDep uses information such as income, employment, educational qualification, and house ownership from census meshblock data to create a deprivation scale ranging from 1 (least deprived 10% of meshblocks) to 10 (most

deprived 10% of meshblocks). NZDep was coded into three deprivation levels in this study: low (NZDep 1–3), medium (NZDep 4–7), and high (NZDep 8–10).

*Area ethnic similarity* was operationalised as the percentage in the participant's neighbourhood, measured in area units, that had the same administratively-prioritised ethnicity as them. An area unit is an aggregation of meshblock units approximately the size of a suburb (median = 2,000 residents; Statistics New Zealand, 2017). Ethnic composition data from the 2013 Census meshblock dataset (Statistics New Zealand, 2014a), which contains broad total response ethnic grouping counts by area unit, were linked to the analytic datasets via meshblock code. To calculate the percentage, the count for the ethnic group which matched participants' administratively-prioritised ethnicity was divided by the total number of responses in that area unit. Skewness and kurtosis statistics were within acceptable bounds of normality (see Table 5.1). To improve comparability and interpretability, the variable was standardised ($M = 0$, $SD = 1$) within each sample before analysis.

### 5.3.3 Data Analysis

For each sample, a main effects multinomial logistic regression model was specified to examine the relationship that individual and contextual demographic characteristics had with discrepancies between administratively-prioritised and self-prioritised ethnicity. The generalised variance inflation factors calculated for each model did not indicate any collinearity issues (all < 2). Then, we explored interaction effects between ethnic combination and each of the applicable demographic characteristics, as well as the interaction between area deprivation and area ethnic similarity. Interaction effects were tested in the same way for each sample: each applicable two-way interaction (e.g., ethnic combination × sex, ethnic combination × age group, etc.) was added separately to the main effects model. An interaction was retained if a likelihood-ratio test indicated that it significantly improved model fit. Where more than one interaction improved model fit, they were added iteratively beginning with the interaction with

the largest likelihood-ratio value. An interaction was retained in the final model if its addition yielded a significant likelihood-ratio test. Results with a *p*-value less than .05 were considered significant.

Events per variable, defined as the frequency of the smallest outcome category divided by the degrees of freedom required by the independent variables, is an important consideration in logistic regression models (de Jong et al., 2019; Vittinghoff & McCulloch, 2007). At least 10 events per variable is a rule-of-thumb to ensure unbiased results, although a simulation study suggests that severe problems are uncommon when this value is above 4 (Vittinghoff & McCulloch, 2007). Events per variable for the child, adolescent, and adult main effects models were 15.47, 3.06, and 8.81, respectively. Therefore, the adolescent model needs to be interpreted with caution. Due to limited events per variable in the adolescent model, interaction effects in adolescents were explored using binary logistic regression, with the smallest outcome category (discrepant—no main ethnic group) omitted. Events per variable for the final child, adolescent, and adult interaction models were 8.00, 29.91, and 5.45, respectively.

All analyses were conducted using R version 4.0.2 (R Core Team, 2020).

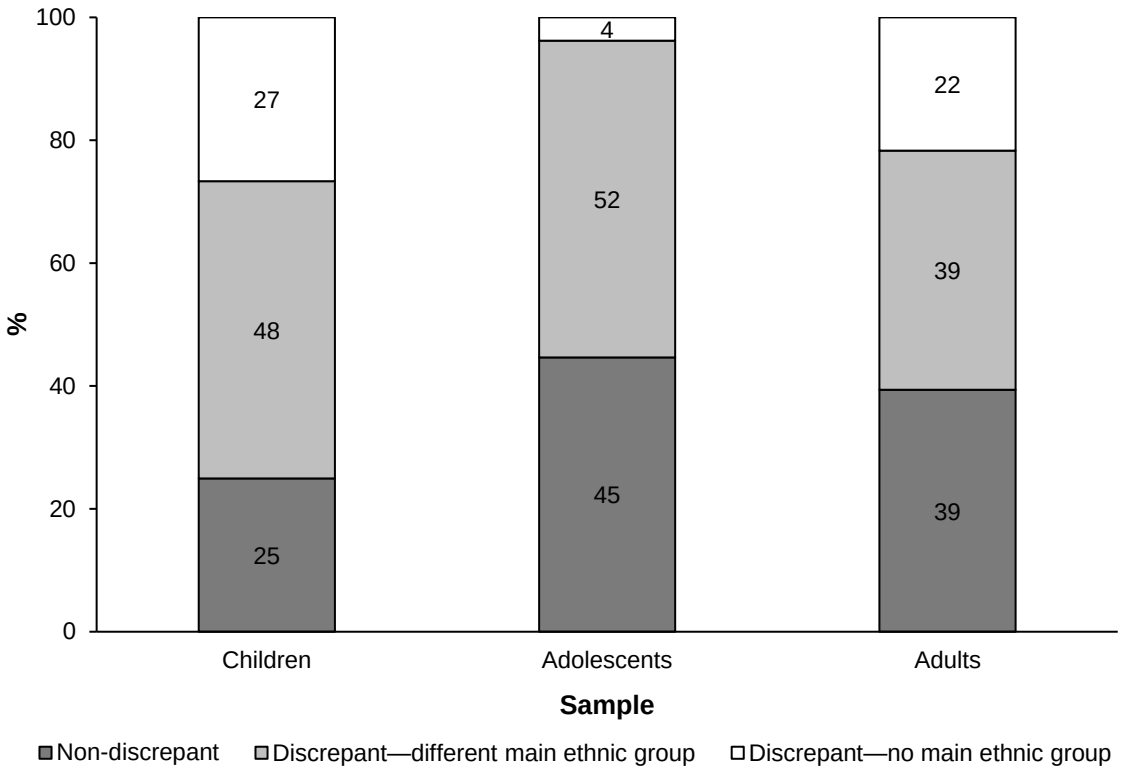
## 5.4 Results

### 5.4.1 Descriptive Statistics

The prevalence of discrepancies between administratively-prioritised and self-prioritised ethnicity for each sample is shown in Figure 5.1. Discrepancies were observed in around 60% of multi-ethnic adolescents and adults, and 75% of multi-ethnic children. Discrepancies due to a different main ethnic group (64% of discrepancies in children and adults, and 93% of discrepancies in adolescents) were more common than discrepancies due to no main ethnic group. Note that, because of differences in survey methodology, the "discrepant—

**Figure 5.1**

*Discrepancies Between Administratively-Prioritised and Self-Prioritised Ethnicity by Sample*




*Note.* Child responses were collected by proxy from the child's mother.

no main ethnic group" category was coded from a single "I cannot choose a main ethnic group" option for adolescents, but was an aggregation of three options for children and adults: two main ethnic groups, don't know, or refused response. There was significant heterogeneity in discrepancy rates by ethnic combination for children and adolescents, but not adults (see Supplementary Tables B.1 to B.3 in Appendix B).

### 5.4.2 Individual Characteristics Associated with Discrepancies

A main effects multinomial logistic regression model was specified for each sample to examine the relationship that individual and contextual demographic characteristics had with discrepancies between administratively-prioritised and self-prioritised ethnicity (see Table 5.2). Compared to Māori/European (the reference group), discrepancies in children and adolescents

due to a different main ethnic group were less common in ethnic combinations which included the Pacific grouping (i.e., Pacific/European, Māori/Pacific, and Māori/Pacific/European), as well as the residual "other combinations" grouping (all odds ratios [*OR*s] < 0.65). Discrepancies due to a different main ethnic group were also less likely to occur in Asian/European adolescents (*OR* = 0.64), but were more likely to occur in Asian/European children (*OR* = 4.15). The occurrence of discrepancy due to a different main ethnic group did not differ significantly between the remaining ethnic combinations and the Māori/European referent. In terms of discrepancies due to no main ethnic group, rates were lower in Pacific/European children (*OR* = 0.61), and higher in Asian/European children (*OR* = 3.77) and adolescents (*OR* = 2.36), when compared to their Māori/European counterparts. No significant differences were observed in other ethnic combinations. In adults, there were no significant ethnic differences in either type of discrepancy (i.e., different main ethnic group or no main ethnic group).

Besides ethnic combination, the only other individual characteristic significantly associated with discrepancies due to a different main ethnic group was birthplace in the adolescent sample. Compared to their counterparts born in Aotearoa New Zealand, adolescents born overseas were significantly less likely to select a main ethnic group discrepant from their administratively-prioritised ethnicity (*OR* = 0.72). However, birthplace was not significantly associated with discrepancies due to no main ethnic group. Where examined, sex, age, and highest educational attainment were not significantly associated with either type of discrepancy in any of the samples (e.g., sex was not examined in adults because all were female).

### 5.4.3 Contextual Characteristics Associated with Discrepancies

In terms of contextual characteristics, area ethnic similarity and socioeconomic deprivation were inversely associated with discrepancies due to a different main ethnic group

**Table 5.2**

*Main Effects Multinomial Logistic Regression Model of Demographic Characteristics Predicting Discrepancies[a] Between Administratively-Prioritised and Self-Prioritised Ethnicity by Sample*

| | Children[b] | | | | Adolescents | | | | Adults | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Different main ethnic group | | No main ethnic group | | Different main ethnic group | | No main ethnic group | | Different main ethnic group | | No main ethnic group | |
| | *B (SE)* | *OR* | *B (SE)* | *OR* | *B (SE)* | *OR* | *B (SE)* | *OR* | *B (SE)* | *OR* | *B (SE)* | *OR* |
| **Parameter estimates** | | | | | | | | | | | | |
| Intercept | 1.22 (0.21)*** | 3.39 | 0.72 (0.23)** | 2.06 | 1.03 (0.14)*** | 2.80 | -1.90 (0.32)*** | 0.15 | 0.48 (0.24)* | 1.62 | -0.12 (0.29) | 0.89 |
| Ethnic combination[c] (ref. Māori/European) | | | | | | | | | | | | |
| Pacific/European | -0.55 (0.20)** | 0.58 | -0.49 (0.23)* | 0.61 | -0.58 (0.13)*** | 0.56 | -0.28 (0.39) | 0.75 | 0.09 (0.24) | 1.10 | 0.45 (0.26) | 1.56 |
| Asian/European | 1.42 (0.35)*** | 4.15 | 1.33 (0.37)*** | 3.77 | -0.45 (0.16)** | 0.64 | 0.86 (0.34)* | 2.36 | - | - | - | - |
| MELAA/European | 0.93 (0.54) | 2.54 | 0.71 (0.59) | 2.03 | -0.30 (0.27) | 0.74 | 0.71 (0.51) | 2.03 | - | - | - | - |
| Māori/Pacific | -0.94 (0.25)*** | 0.39 | -0.12 (0.24) | 0.89 | -0.99 (0.25)*** | 0.37 | -0.49 (0.76) | 0.61 | -0.15 (0.30) | 0.86 | -0.38 (0.34) | 0.69 |
| Pacific/Asian | - | - | - | - | 0.39 (0.24) | 1.48 | 0.81 (0.56) | 2.25 | - | - | - | - |
| Asian/Other | -0.46 (0.30) | 0.63 | -0.41 (0.32) | 0.66 | - | - | - | - | - | - | - | - |
| Māori/Pacific/European | -0.44 (0.21)* | 0.64 | -0.10 (0.22) | 0.91 | -0.52 (0.19)** | 0.60 | -0.19 (0.55) | 0.83 | 0.29 (0.31) | 1.33 | 0.01 (0.37) | 1.01 |
| Other combinations | -0.56 (0.19)** | 0.57 | -0.23 (0.21) | 0.79 | -1.02 (0.16)*** | 0.36 | 0.13 (0.37) | 1.14 | 0.28 (0.24) | 1.32 | 0.22 (0.28) | 1.25 |
| Sex (ref. female) | | | | | | | | | | | | |
| Male | -0.05 (0.12) | 0.95 | -0.08 (0.13) | 0.93 | 0.04 (0.09) | 1.04 | -0.37 (0.23) | 0.69 | - | - | - | - |
| Age group (years) | | | | | | | | | | | | |
| Children (all 4–5) | - | - | - | - | - | - | - | - | - | - | - | - |
| Adolescents (ref. <16) | | | | | | | | | | | | |
| ≥16 | - | - | - | - | -0.08 (0.10) | 0.92 | -0.24 (0.24) | 0.79 | - | - | - | - |
| Adults (ref. <30) | | | | | | | | | | | | |
| ≥30 | - | - | - | - | - | - | - | - | -0.15 (0.16) | 0.86 | -0.19 (0.19) | 0.83 |
| Birthplace (ref. Aotearoa New Zealand) | | | | | | | | | | | | |
| Overseas | - | - | - | - | -0.33 (0.13)* | 0.72 | 0.21 (0.28) | 1.23 | -0.17 (0.26) | 0.84 | -0.60 (0.32) | 0.55 |
| Highest educational attainment[d] (ref. secondary school qualification or below) | | | | | | | | | | | | |
| Certificate or diploma | 0.28 (0.14) | 1.33 | 0.03 (0.16) | 1.03 | - | - | - | - | 0.03 (0.17) | 1.03 | -0.30 (0.19) | 0.74 |
| Bachelor's or above | 0.13 (0.16) | 1.14 | -0.11 (0.18) | 0.89 | - | - | - | - | -0.09 (0.20) | 0.91 | -0.35 (0.24) | 0.70 |
| Urbanicity (ref. main urban) | | | | | | | | | | | | |
| Non-main urban | 0.65 (0.22)** | 1.92 | 0.66 (0.23)** | 1.93 | 0.03 (0.15) | 1.03 | 0.25 (0.37) | 1.28 | 0.44 (0.23) | 1.56 | -0.10 (0.28) | 0.90 |
| Rural | 0.11 (0.24) | 1.11 | -0.34 (0.29) | 0.71 | -0.11 (0.14) | 0.90 | 0.04 (0.37) | 1.04 | 0.48 (0.30) | 1.61 | 0.03 (0.38) | 1.03 |
| Socioeconomic deprivation (ref. low dep.) | | | | | | | | | | | | |
| Medium dep. | -0.43 (0.19)* | 0.65 | -0.36 (0.21) | 0.69 | -0.49 (0.14)*** | 0.61 | -0.62 (0.29)* | 0.54 | -0.41 (0.22) | 0.66 | -0.20 (0.27) | 0.82 |
| High dep. | -0.85 (0.20)*** | 0.43 | -0.86 (0.22)*** | 0.42 | -0.67 (0.15)*** | 0.51 | -0.77 (0.33)* | 0.46 | -0.87 (0.23)*** | 0.42 | -0.26 (0.29) | 0.77 |
| Area ethnic similarity[e] | -0.43 (0.07)*** | 0.65 | -0.13 (0.07) | 0.88 | -0.45 (0.06)*** | 0.63 | -0.41 (0.15)** | 0.66 | -0.28 (0.08)*** | 0.76 | -0.11 (0.09) | 0.90 |
| **Model summary** | | | | | | | | | | | | |
| Model $\chi^2$ (*df*) | 248.52 (30)*** | | | | 327.70 (30)*** | | | | 73.96 (26)*** | | | |
| Pseudo-$R^2$ | | | | | | | | | | | | |
| McFadden | 0.06 | | | | 0.08 | | | | 0.03 | | | |
| Cox & Snell | 0.13 | | | | 0.13 | | | | 0.07 | | | |
| Nagelkerke | 0.14 | | | | 0.16 | | | | 0.08 | | | |

*Note.* A hyphen indicates that the variable was not applicable for the sample. MELAA = Middle Eastern, Latin American, and African. [a]Levels were: non-discrepant (0), discrepant—different main ethnic group (1), and discrepant—no main ethnic group (2). [b]Child responses were collected by proxy from the child's mother. [c]Combinations with $n < 50$ were aggregated into the "other combinations" category. [d]Mother's highest educational attainment used for children. [e]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; standardised ($M = 0$, $SD = 1$).
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

in each sample (see Table 5.2). This results in *OR*s under 1, which are less intuitive to interpret. To aid interpretability, the text below uses reversed *OR*s (e.g., high deprivation:low deprivation was converted to low deprivation:high deprivation by taking the inverse of the *OR* [i.e., $1/e^b$]).

When all other demographic characteristics were held constant, discrepancies due to a different main ethnic group were more likely to occur in participants who lived in a neighbourhood where there was a lower percentage of people who had the same administratively-prioritised ethnicity as them (*OR*s between 1.32 and 1.58 per *SD* difference in ethnic composition). For example, a Māori/European participant who lived in a neighbourhood with lower Māori density (their administratively-prioritised ethnicity) was less likely to self-prioritise their Māori identity than a Māori/European participant who lived in a neighbourhood with higher Māori density. Likewise, discrepancies due to a different main ethnic group were more likely to occur in participants who lived in a less socioeconomically-deprived area (*OR*s between high and low deprivation ranged between 1.96 and 2.38). For example, a Māori/European participant living in a low deprivation area was more likely to self-prioritise their European identity, whereas a Māori/European participant living in a high deprivation area was more likely to self-prioritise their Māori identity.

Among adolescents, area ethnic similarity (*OR* = 1.51) and socioeconomic deprivation (*OR* = 2.15) were also inversely associated with discrepancies due to no main ethnic group. In other words, adolescents were more likely to select “I can’t choose only one ethnic group” if they lived in an area with a lower percentage of people who had the same administratively-prioritised ethnicity, or a less deprived area. Among children, deprivation was also inversely associated with discrepancies due to no main ethnic group (*OR* = 2.36), and the inverse relationship between area ethnic similarity and discrepancies due to no main ethnic group was approaching significance (*OR* = 1.14, $p = .077$). Among adults, neither area ethnic similarity

nor socioeconomic deprivation was significantly associated with discrepancies due to no main ethnic group.

For children, urbanicity was significantly associated with both discrepancies due to a different main ethnic group and discrepancies due to no main ethnic group, but no significant association was observed for adolescents and adults. Children living in non-main urban areas were more likely to have each type of discrepancy (*OR* = 1.91 and 1.93, respectively) compared to their counterparts living in main urban areas.

### 5.4.4 Interaction Effects Associated with Discrepancies

Next, we explored whether there were significant interaction effects in each sample between ethnic combination and each of the applicable demographic characteristics, as well as between area ethnic similarity and socioeconomic deprivation. For children and adults, likelihood-ratio tests showed that addition of an interaction between ethnic combination and highest educational attainment (of mothers for children) significantly improved model fit over their respective main effects models (children: $\chi^2(28) = 44.36$, $p = .026$; adults: $\chi^2(16) = 34.32$, $p = .005$). None of the other interactions tested significantly improved model fit for children and adults after this interaction was included.

The interaction terms between ethnic combination and mother's highest educational attainment in the child model showed that, for most ethnic combinations, education was not significantly associated with either type of discrepancy tested (i.e., discrepancy due to a different main ethnic group, or discrepancy due to no main ethnic group; see Table 5.3). However, Asian/European children whose mothers' highest qualification was at secondary school level or below had a greater likelihood of both types of discrepancy compared to Asian/European children whose mothers held higher qualifications (i.e., certificate-level or above). No mother in the former qualification group (0 from 31) selected a self-prioritised

**Table 5.3**

*Interaction Effects Multinomial Logistic Regression Model of Demographic Characteristics Predicting Discrepancies[a] Between Administratively-Prioritised and Self-Prioritised Ethnicity for Children[b]*

| | Different main ethnic group | | No main ethnic group | |
|---|---|---|---|---|
| | *B (SE)* | *OR* | *B (SE)* | *OR* |
| **Parameter estimates** | | | | |
| Intercept | 1.17 (0.23)*** | 3.21 | 0.79 (0.25)** | 2.20 |
| Ethnic combination[c] (ref. Māori/European) | | | | |
| Pacific/European | -0.65 (0.33)* | 0.52 | -1.14 (0.40)** | 0.32 |
| Asian/European[d] | 13.44 (0.42)*** | 689945.69 | 13.22 (0.42)*** | 549598.75 |
| MELAA/European | 1.48 (1.05) | 4.40 | 0.95 (1.14) | 2.59 |
| Māori/Pacific | -0.83 (0.36)* | 0.43 | 0.09 (0.33) | 1.10 |
| Asian/Other | -0.90 (0.65) | 0.41 | -0.68 (0.64) | 0.50 |
| Māori/Pacific/European | -0.11 (0.31) | 0.90 | -0.14 (0.33) | 0.87 |
| Other combinations | -0.49 (0.31) | 0.61 | -0.29 (0.32) | 0.75 |
| Sex (ref. female) | | | | |
| Male | -0.06 (0.12) | 0.94 | -0.05 (0.13) | 0.95 |
| Mother's highest educational attainment (ref. secondary school qualification or below) | | | | |
| Certificate or diploma | 0.27 (0.22) | 1.31 | -0.10 (0.25) | 0.90 |
| Bachelor's or above | 0.37 (0.24) | 1.45 | -0.10 (0.27) | 0.91 |
| Urbanicity (ref. main urban) | | | | |
| Non-main urban | 0.64 (0.22)** | 1.90 | 0.66 (0.23)** | 1.93 |
| Rural | 0.09 (0.24) | 1.10 | -0.38 (0.29) | 0.68 |
| Socioeconomic deprivation (ref. low dep.) | | | | |
| Medium dep. | -0.43 (0.19)* | 0.65 | -0.40 (0.21) | 0.67 |
| High dep. | -0.86 (0.20)*** | 0.42 | -0.91 (0.22)*** | 0.40 |
| Area ethnic similarity[e] | -0.42 (0.07)*** | 0.65 | -0.12 (0.07) | 0.89 |
| Mother's highest educational attainment × ethnic combination[c] | | | | |
| Certificate or diploma × Pacific/European | 0.64 (0.47) | 1.91 | 0.98 (0.55) | 2.67 |
| Certificate or diploma × Asian/European[d] | -11.05 (0.74)*** | 0.00 | -11.18 (0.77)*** | 0.00 |
| Certificate or diploma × MELAA/European | -0.95 (1.51) | 0.39 | -0.17 (1.63) | 0.84 |
| Certificate or diploma × Māori/Pacific | 0.30 (0.51) | 1.34 | -0.20 (0.52) | 0.82 |
| Certificate or diploma × Asian/Other | 0.68 (0.82) | 1.97 | -0.12 (0.91) | 0.88 |
| Certificate or diploma × Māori/Pacific/European | -0.52 (0.45) | 0.59 | 0.11 (0.49) | 1.12 |
| Certificate or diploma × Other combinations | -0.35 (0.45) | 0.70 | 0.30 (0.46) | 1.36 |
| Bachelor's or above × Pacific/European | -0.46 (0.51) | 0.63 | 0.93 (0.56) | 2.54 |
| Bachelor's or above × Asian/European | -12.76 (0.49)*** | 0.00 | -12.38 (0.51)*** | 0.00 |
| Bachelor's or above × MELAA/European | -0.83 (1.30) | 0.44 | -0.46 (1.41) | 0.63 |
| Bachelor's or above × Māori/Pacific | -2.38 (1.14)* | 0.09 | -1.26 (0.78) | 0.28 |
| Bachelor's or above × Asian/Other | 0.31 (0.78) | 1.37 | 0.51 (0.78) | 1.67 |
| Bachelor's or above × Māori/Pacific/European | -0.59 (0.56) | 0.55 | 0.06 (0.60) | 1.06 |
| Bachelor's or above × Other combinations | 0.15 (0.48) | 1.17 | -0.27 (0.55) | 0.76 |
| **Model summary** | | | | |
| Model $\chi^2$ (*df*) | 292.88 (58)*** | | | |
| Pseudo-$R^2$ | | | | |
| McFadden | 0.07 | | | |
| Cox & Snell | 0.15 | | | |
| Nagelkerke | 0.17 | | | |

*Note.* MELAA = Middle Eastern, Latin American, and African.

[a]Levels were: non-discrepant (0), discrepant—different main ethnic group (1), and discrepant—no main ethnic group (2).

[b]Child responses were collected by proxy from the child's mother.

[c]Combinations with $n < 50$ were aggregated into the "other combinations" category.

[d]Parameter estimates for the Asian/European terms were unusually extreme because no mother in the reference educational attainment category selected a self-prioritised ethnicity for their child that matched administrative-prioritisation.

[e]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; standardised ($M = 0$, $SD = 1$).

*$p < .05$. **$p < .01$. ***$p < .001$.

**Table 5.4**

*Interaction Effects Multinomial Logistic Regression Model of Demographic Characteristics Predicting Discrepancies[a] Between Administratively-Prioritised and Self-Prioritised Ethnicity for Adults*

| | Different main ethnic group | | No main ethnic group | |
|---|---|---|---|---|
| | *B (SE)* | *OR* | *B (SE)* | *OR* |
| **Parameter estimates** | | | | |
| Intercept | 0.50 (0.25)* | 1.65 | -0.18 (0.30) | 0.83 |
| Ethnic combination[b] (ref. Māori/European) | | | | |
| Pacific/European | 0.11 (0.39) | 1.11 | 0.25 (0.41) | 1.28 |
| Māori/Pacific | -0.04 (0.39) | 0.96 | -0.56 (0.48) | 0.57 |
| Māori/Pacific/European | 0.35 (0.48) | 1.41 | 0.28 (0.51) | 1.32 |
| Other combinations | -0.16 (0.44) | 0.85 | 0.52 (0.41) | 1.68 |
| Age group (years; ref. <30) | | | | |
| ≥30 | -0.17 (0.16) | 0.84 | -0.18 (0.19) | 0.84 |
| Birthplace (ref. Aotearoa New Zealand) | | | | |
| Overseas | -0.19 (0.27) | 0.82 | -0.61 (0.33) | 0.54 |
| Highest educational attainment (ref. secondary school qualification or below) | | | | |
| Certificate or diploma | -0.05 (0.22) | 0.95 | -0.13 (0.25) | 0.88 |
| Bachelor's or above | -0.07 (0.24) | 0.93 | -0.55 (0.30) | 0.58 |
| Urbanicity (ref. main urban) | | | | |
| Non-main urban | 0.41 (0.23) | 1.51 | -0.12 (0.28) | 0.88 |
| Rural | 0.46 (0.30) | 1.58 | 0.06 (0.38) | 1.06 |
| Socioeconomic deprivation (ref. low dep.) | | | | |
| Medium dep. | -0.38 (0.22) | 0.69 | -0.15 (0.28) | 0.86 |
| High dep. | -0.86 (0.24)*** | 0.42 | -0.21 (0.29) | 0.81 |
| Area ethnic similarity[c] | -0.29 (0.08)*** | 0.75 | -0.10 (0.09) | 0.91 |
| Highest educational attainment × ethnic combination[b] | | | | |
| Certificate or diploma × Pacific/European | 0.16 (0.53) | 1.17 | 0.13 (0.56) | 1.14 |
| Certificate or diploma × Māori/Pacific | 0.01 (0.64) | 1.01 | 0.55 (0.72) | 1.73 |
| Certificate or diploma × Māori/Pacific/European | 0.44 (0.66) | 1.55 | -1.89 (1.19) | 0.15 |
| Certificate or diploma × Other combinations | 0.33 (0.55) | 1.39 | -1.31 (0.67)* | 0.27 |
| Bachelor's or above × Pacific/European | -0.37 (0.65) | 0.69 | 0.63 (0.68) | 1.87 |
| Bachelor's or above × Māori/Pacific | -0.94 (0.93) | 0.39 | -0.30 (1.21) | 0.74 |
| Bachelor's or above × Māori/Pacific/European[d] | -13.54 (308.77) | 0.00 | 0.50 (0.89) | 1.65 |
| Bachelor's or above × Other combinations | 0.76 (0.55) | 2.13 | 0.23 (0.61) | 1.26 |
| **Model summary** | | | | |
| Model $\chi^2$ (*df*) | 108.28 (42)*** | | | |
| Pseudo-$R^2$ | | | | |
| McFadden | 0.05 | | | |
| Cox & Snell | 0.10 | | | |
| Nagelkerke | 0.11 | | | |

*Note.* MELAA = Middle Eastern, Latin American, and African.

[a]Levels were: non-discrepant (0), discrepant—different main ethnic group (1), and discrepant—no main ethnic group (2).

[b]Combinations with $n < 50$ were aggregated into the "other combinations" category.

[c]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; standardised ($M = 0$, $SD = 1$).

[d]Unusually extreme parameter estimate as no Māori/Pacific/European participant with a Bachelor's or above was in the "discrepant—different main ethnic group" category.

*$p < .05$. **$p < .01$. ***$p < .001$.

ethnicity for their child that matched administrative-prioritisation (Asian), so the parameter estimates for the Asian/European terms tended to be unusually extreme. For adults, highest educational attainment was also not significantly related to either type of discrepancy for most ethnic combinations (see Table 5.4). However, for the residual "other combinations" grouping, discrepancy rates due to no main ethnic group was marginally lower for those who had certificates or diplomas compared to those who only had a secondary school qualification or below ($p = .049$).

For adolescents, binary logistic regression was conducted to examine interaction effects associated with discrepancies due to a different main ethnic group (discrepancies due to no main ethnic group was not examined due to low frequency). Likelihood-ratio tests showed that ethnic combination × birthplace significantly improved the main effects model conducted with binary outcomes ($\chi^2(7) = 31.16$, $p < .001$), and subsequent addition of ethnic combination × socioeconomic deprivation further improved model fit ($\chi^2(14) = 30.45$, $p = .007$). No other tested interaction significantly improved this model.

For adolescents who identified with a dual combination involving the European grouping (Māori/European, Pacific/European, Asian/European, and MELAA/European), discrepancy rates due to a different main ethnic group were lower if they were born overseas rather than in Aotearoa New Zealand (i.e., those born overseas were more likely to self-prioritise their Indigenous or ethnic minority group; see Table 5.5). Contrastingly, discrepancy rates in Pacific/Asian and Māori/Pacific/European adolescents were higher in those born overseas (i.e., they were more likely to self-prioritise Asian and Pacific, respectively). Birthplace associations could not be accurately ascertained for Māori/Pacific because nearly all were born in Aotearoa New Zealand (hence unusually large parameter estimates). In terms of socioeconomic deprivation, lower deprivation was associated with increased discrepancies due to a different main ethnic group for all ethnic combinations except Māori/Pacific. In

**Table 5.5**

*Interaction Effects Binary Logistic Regression Model of Demographic Characteristics Predicting Discrepancies[a] Between Administratively-Prioritised and Self-Prioritised Ethnicity for Adolescents*

| | Different main ethnic group | |
|---|---|---|
| | *B (SE)* | *OR* |
| **Parameter estimates** | | |
| Intercept | 1.07 (0.19)*** | 2.92 |
| Ethnic combination[b] (ref. Māori/European) | | |
| Pacific/European | -0.46 (0.43) | 0.63 |
| Asian/European | -0.27 (0.32) | 0.76 |
| MELAA/European | 0.54 (0.73) | 1.71 |
| Māori/Pacific | -2.77 (1.13)* | 0.06 |
| Pacific/Asian | -0.96 (0.88) | 0.38 |
| Māori/Pacific/European | -0.51 (0.51) | 0.60 |
| Other combinations | -0.90 (0.41)* | 0.41 |
| Sex (ref. female) | 0.05 (0.09) | |
| Male | | 1.05 |
| Age group (years; ref. <16) | -0.09 (0.10) | |
| ≥16 | | 0.91 |
| Birthplace (ref. Aotearoa New Zealand) | | |
| Overseas | -0.78 (0.37)* | 0.46 |
| Urbanicity (ref. main urban) | | |
| Non-main urban | 0.04 (0.15) | 1.04 |
| Rural | -0.13 (0.15) | 0.87 |
| Socioeconomic deprivation (ref. low dep.) | | |
| Medium dep. | -0.45 (0.20)* | 0.64 |
| High dep. | -0.76 (0.21)*** | 0.47 |
| Area ethnic similarity[c] | -0.44 (0.06)*** | 0.65 |
| Birthplace × ethnic combination[b] | | |
| Overseas × Pacific/European | 0.44 (0.46) | 1.55 |
| Overseas × Asian/European | -0.22 (0.48) | 0.80 |
| Overseas × MELAA/European | -0.50 (0.74) | 0.60 |
| Overseas × Māori/Pacific[d] | 13.23 (324.01) | 559222.80 |
| Overseas × Pacific/Asian | 2.05 (0.58)*** | 7.79 |
| Overseas × Māori/Pacific/European | 3.03 (1.16)** | 20.70 |
| Overseas × Other combinations | 0.47 (0.46) | 1.60 |
| Socioeconomic deprivation × ethnic combination[b] | | |
| Medium dep. × Pacific/European | -0.06 (0.48) | 0.94 |
| Medium dep. × Asian/European | -0.03 (0.39) | 0.97 |
| Medium dep. × MELAA/European | -0.22 (0.60) | 0.81 |
| Medium dep. × Māori/Pacific | -0.78 (1.53) | 0.46 |
| Medium dep. × Pacific/Asian | 0.38 (0.96) | 1.46 |
| Medium dep. × Māori/Pacific/European | -0.86 (0.62) | 0.42 |
| Medium dep. × Other combinations | 0.01 (0.46) | 1.01 |
| High dep. × Pacific/European | -0.16 (0.46) | 0.85 |
| High dep. × Asian/European | 0.07 (0.45) | 1.07 |
| High dep. × MELAA/European | 0.06 (0.77) | 1.07 |
| High dep. × Māori/Pacific | 2.41 (1.17)* | 11.13 |
| High dep. × Pacific/Asian | 0.54 (0.90) | 1.72 |
| High dep. × Māori/Pacific/European | 0.27 (0.58) | 1.31 |
| High dep. × Other combinations | -0.34 (0.45) | 0.71 |
| **Model summary** | | |
| Model $\chi^2$ (*df*) | 355.50 (36)*** | |
| Pseudo-$R^2$ | | |
| McFadden | 0.11 | |
| Cox & Snell | 0.14 | |
| Nagelkerke | 0.19 | |

*Note.* MELAA = Middle Eastern, Latin American, and African.

[a]Levels were: non-discrepant (0), and discrepant—different main ethnic group (1). Discrepant—no main ethnic group (2) was excluded due to small subgroup size ($n$ = 92).

[b]Combinations with $n$ < 50 were aggregated into the "other combinations" category.

[c]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; standardised ($M$ = 0, $SD$ = 1).

[d]Unusually extreme parameter estimate as nearly all Māori/Pacific adolescents were born in Aotearoa New Zealand.

*$p$ < .05. **$p$ < .01. ***$p$ < .001.

Māori/Pacific adolescents, the relationship was reversed—higher deprivation was associated with higher discrepancy rates (i.e., self-prioritisation of Pacific over Māori).

## 5.5 Discussion

The current study utilised large datasets from three age cohorts to investigate individual and contextual demographic characteristics associated with discrepancies between two methods commonly used to prioritise multiple ethnic identifications: administrative-prioritisation and self-prioritisation. Overall, ethnic group membership as classified by the two prioritisation methods differed for over half of multi-ethnic adolescents and adults, and up to three-quarters of multi-ethnic children, indicating that the methods are not interchangeable. Discrepancies due to participants self-prioritising a different main ethnic group from that outputted by the administrative-prioritisation hierarchy were more common than discrepancies due to participants not selecting a main ethnic group. Differences in survey response options may have contributed to a relatively higher proportion of discrepancy due to no main ethnic group in children and adults (options included selecting two main ethnic groups, "don't know", and refuse to respond) compared to adolescents (single option of "I can't choose only one ethnic group"; Yao et al., 2021). Although we do not have information on why participants chose not to select a main ethnicity, previous research indicates this decision is likely to reflect a declaration of multi-ethnic identity or a rejection of traditional ethnic categories (Aspinall & Song, 2013).

Even after accounting for demographic characteristics, discrepancies due to a different main ethnic group were less likely to occur in child and adolescent ethnic combinations which included the Pacific grouping (e.g., Pacific/European, Māori/Pacific, and Māori/Pacific/European), and discrepancies due to no main ethnic group were more likely to occur in Asian/European children and adolescents. This importantly indicates that researchers'

choice of prioritisation method affects ethnic combinations differently. Further qualitative research is needed to explore the specific reasons underlying self-prioritisation responses of young people in each ethnic combination. However, the heterogeneity by ethnic combination among the younger age groups may reflect increasing flexibility for young people of some ethnic combinations to identify their main ethnic group in a way that matches their private self-definitions, rather than adhere to societal prescriptions. The finding that Asian/European young people were more likely not to select a main ethnic group may reflect reluctance to self-prioritise as Asian due to higher rates of ethnicity-related bullying experienced by Asian adolescents (Crengle et al., 2012), in conjunction with difficulty passing as European due to the stereotype that those with Asian heritage are "perpetual foreigners" (S. J. Lee et al., 2009).

Apart from ethnic combination, contextual characteristics generally had a stronger relationship with discrepancies between administratively-prioritised and self-prioritised ethnicity than individual characteristics. As discussed in the following two sections, ethnic groups' neighbourhood ethnic composition and socioeconomic deprivation profiles were systematically associated with prioritisation method in each age cohort examined, suggesting that similar underlying psychological and sociological processes undergird each developmental stage. Conversely, most individual demographic characteristics examined (i.e., sex, age, and highest educational attainment) were not associated with discrepancies. Birthplace for adolescents was the only exception: Māori/European, Pacific/European, Asian/European, and MELAA/European adolescents born in Aotearoa New Zealand were more likely than their counterparts born overseas to self-prioritise a European ethnic group; and Pacific/Asian and Māori/Pacific/European adolescents born overseas were more likely than their counterparts born in Aotearoa New Zealand to self-prioritise an Asian and Pacific ethnic group, respectively.

To illustrate what these results mean for ethnic group composition, consider the ethnic classification of adolescents who identify with a dual combination comprising an Indigenous

or ethnic minority group (Māori, Pacific, Asian, or MELAA; hereafter collectively referred to as "minority groups" for brevity) and the ethnic majority group (European)—the most common combination in Aotearoa New Zealand. Under administrative-prioritisation, these combinations will always be classified to its respective minority group. However, under self-prioritisation, multi-ethnic individuals were more likely to be classified to an ethnic minority group if they were living in neighbourhoods of higher ethnic minority density or higher socioeconomic deprivation, or were born overseas. These systematic discrepancies highlight the association between sociocultural factors and main ethnic group response. Consequentially, researchers' choice of prioritisation method can bias ethnic group composition and resulting analyses.

### 5.5.1 Associations with Neighbourhood Ethnic Composition

For all age groups, living in a neighbourhood that had a higher density of the participant's administratively-prioritised ethnicity was associated with self-prioritisation of their administratively-prioritised ethnicity (i.e., lower discrepancy due to a different main ethnic group). In other words, self-prioritised ethnic groups included participants from more ethnically *homo*geneous communities, whereas administratively-prioritised ethnic groups included participants from more ethnically *hetero*geneous communities. This is consistent with U.S. research, which generally found that multiracial individuals, regardless of age, were more likely to select the race that was more prevalent in their neighbourhood (Bratter, 2007; Campbell, 2007; D. R. Harris & Sim, 2002; Herman, 2004; Roth, 2005; Schenker & Parker, 2003). There has been limited research conducted in this area in Aotearoa New Zealand, but Kukutai's (2004) study of Māori/European women also found that those who lived in areas of higher Māori density were more likely to self-prioritise their Māori ethnic group. Such contexts are more likely to be affirming of Māori identity, leading to higher rates of self-prioritisation as Māori (Bécares et al., 2013; Kukutai, 2004). For example, Bécares et al.'s (2013) study using

the Māori sub-sample of the cross-sectional 2006/2007 New Zealand Health Survey of adults (aged 15 years and over) found that higher Māori ethnic density was associated with reduced racial discrimination experiences after area deprivation was controlled for.

Indeed, researchers argue that living in a neighbourhood with higher density of a particular ethnic group both increases exposure and sense of belonging to that ethnicity, which act as "pull" factors towards identification with that ethnic group; and decreases negative experiences such as racial discrimination, which act as "push" factors against identification with that ethnic group (Khanna, 2012; Roth, 2005). Ethnic exposure and familiarity may be particularly influential for adolescents given that they are in a key phase of identity development (Khanna, 2012). Because they typically do not select their area of residence, the findings also suggest that adolescents will choose to "fit in" by self-prioritising an identity that matches their social context. For adults, however, the findings could equally reflect self-selection into a neighbourhood that was more congruent with the culture they feel more attached to and/or are able to afford.

In addition, the study suggests that children (by mother proxy) and adolescents living in a neighbourhood with higher density of their administratively-prioritised ethnicity were more likely to self-prioritise that ethnicity as opposed to not selecting a main ethnic group (i.e., lower discrepancy due to no main ethnic group). However, this result is tentative because the relationship was only approaching significance for children, and only a small number of adolescents did not select a main ethnic group. We were unable to explore the association that neighbourhood prevalence of multi-ethnic identification had on non-selection of a main ethnic group due to collinearity concerns with administratively-prioritised ethnic composition in our analytic samples.

### 5.5.2 Associations with Socioeconomic Deprivation

Multi-ethnic children (responded via mother proxy), adolescents, and adults also tended to self-prioritise a main ethnicity that was discrepant from their administratively-prioritised ethnicity (i.e., higher discrepancy due to a different main ethnic group) if they lived in a neighbourhood of lower socioeconomic deprivation. There was no evidence to suggest that the direction of this relationship differed for any ethnic combination except for Māori/Pacific adolescents, for whom there were higher discrepancy rates due to a different main ethnic group (i.e., higher likelihood of self-prioritising Pacific) if they lived in areas of higher deprivation. It is helpful to again interpret these findings in the context of dual ethnic combinations comprising a minority and majority group. Given that minority groups are prioritised ahead of European in the administrative-prioritisation hierarchy, these combinations as classified by administrative-prioritisation result in minority groups that have a lower socioeconomic deprivation profile than minority groups classified by self-prioritisation.

In other words, consistent with previous research in Aotearoa New Zealand and the United States (e.g., Brunsma, 2005; Herman, 2004; Kukutai, 2004; Roth, 2005), multi-ethnic children (via mother report), adolescents, and adults who live in areas of higher socioeconomic deprivation were all more likely to self-prioritise their ethnic minority group. This is likely influenced by the persisting systemic relationship in Aotearoa New Zealand between ethnic minority status and social disadvantage, particularly for Māori and Pacific Peoples, and highlights the interaction between ethnicity and social stratification in the country (Ministry of Social Development, 2016). For example, due to power imbalance and institutional racism, those who identify more strongly as Māori and/or Pacific are more likely to experience labour market discrimination and reduced access to resources, leading to a cycle of socioeconomic disadvantage (Kukutai, 2004). Associated negative societal stereotypes such as "Māori and Pacific Peoples must be disadvantaged" may also structurally constrain the perceived freedom

of main ethnic group choice for these multi-ethnic individuals. For instance, an interview study with Māori adolescents in Aotearoa New Zealand found that many participants felt material disadvantage was a defining characteristic of "being Māori", and one participant shared that he hid his family's wealth from his peers to maintain his Māori identity (Borrell, 2005).

Children (by mother proxy) and adolescents in the current study were also more likely to not select a single main ethnic group than one that matched their administratively-prioritised ethnicity (i.e., higher discrepancy due to no main ethnic group) if they lived in less socioeconomically deprived areas. A similar pattern, using measures such as family income and parental education level, has been observed in parents' identification of children in the United States (Bratter, 2007; Brunsma, 2005; Roth, 2005). However, in the current study, there was no relationship between parents' educational level and whether they selected one main ethnic group for their child, perhaps because a broader index of neighbourhood socioeconomic deprivation, which incorporated educational qualifications, was used (Atkinson et al., 2014). For adults in the current study, neither lower neighbourhood socioeconomic deprivation nor higher educational qualifications were associated with opting to not select a main ethnic group. The significant association of socioeconomic deprivation for children and adolescents, but not adults, suggests a recent shift in more affluent communities towards challenging the notion that multi-ethnic individuals should identify with one main ethnic group alone.

### 5.5.3 Limitations

The relatively consistent association that neighbourhood ethnic composition and socioeconomic deprivation had with discrepancies between administratively-prioritised and self-prioritised ethnicity provide confidence over the associations observed. However, there is a possibility that, for adolescents, the small number who did not select a main ethnic group (potentially due to survey methodology) may lead to biased results for this less prevalent outcome (de Jong et al., 2019). In addition, while the child and adolescent datasets from which

the analytic samples were drawn were broadly nationally representative (Clark et al., 2013; Morton et al., 2013), the adult dataset comprised solely of children's mothers and thus was not representative of the wider adult population in Aotearoa New Zealand. Overlap between the child and adult analytic samples may be an alternate explanation for the similarities observed between child and adult results, but inclusion of the adolescent sample alleviates this concern, particularly because the overall results for children and adolescents (independent samples) were arguably more similar than between children and adults (semi-dependent samples).

Limitations in the measurement and classification of ethnicity also need to be noted. First, ethnic data were collected cross-sectionally in a single context, so fluidity in self-prioritisation could not be examined (e.g., different responses between home and school, or over time; D. R. Harris & Sim, 2002). Second, the current study classified participants' ethnic identifications into broad ethnic groupings. While this is a common way of outputting ethnic data (Statistics New Zealand, 2005), it creates some distance from participants' original responses. Third, we had to aggregate small ethnic combinations due to concerns over statistical power. Incidentally, this limitation illustrates why researchers prefer classification methods that output multiple ethnic identifications into fewer numbers of mutually exclusive categories (Callister et al., 2007; Mays et al., 2003), and hence why we focus on administrative-prioritisation and self-prioritisation in this study.

Finally, while the purpose of this research was to explore demographic characteristics associated with discrepancies between administrative-prioritisation and self-prioritisation, there are other unexamined factors that are also likely to be associated with ethnic identification. Examples include ethnic identity, language proficiency, socially-assigned ethnicity, perceived racism, and experiences of racial discrimination. For instance, the literature suggests multi-ethnic individuals are more likely to self-prioritise an ethnic minority group if they have a stronger sense of ethnic identity in it and/or are perceived by others as belonging to the group

(e.g., due to their physical appearance; Herman, 2004; Roth, 2016). Conversely, they are less likely to self-prioritise an ethnic minority group if they have experienced racism as a result of their ethnic minority status (Callister et al., 2007; Herman, 2004).

### 5.5.4 Implications

Nevertheless, the systematic association that prioritisation method had with the neighbourhood ethnic composition and socioeconomic deprivation profiles of outputted ethnic groups clearly demonstrates the importance of researcher decisions. This naturally leads to the question of which prioritisation method should be used. We concur with previous literature that argues this depends on the research question under investigation (Atatoa Carr et al., 2017; Cormack & Robson, 2010; Yao et al., 2021). For example, if it is important to capture affiliation with certain ethnic groups of interest (e.g., for policy or funding), administrative-prioritisation may be more suitable; whereas if strength of self-affiliation is important, self-prioritisation may be more appropriate. However, the current study has raised additional empirical and ethical implications researchers need to be aware of.

Empirically, the association between ethnic self-prioritisation and socioeconomic deprivation found in this study suggests that the relative estimated associations of ethnicity and deprivation on health and social outcomes may differ depending on researchers' choice of prioritisation method. For instance, the observed association may be larger for ethnicity, and smaller for deprivation, if administrative-prioritisation was used instead of self-prioritisation. Researchers need to be cognisant of the implications this may have on scientific knowledge construction and policymaking, as well as the risk of perpetuating stereotypes and inequities along ethnic and/or deprivation lines.

Ethically, self-prioritisation may initially appear more appropriate because it allows multi-ethnic individuals to select their main ethnic group, whereas administrative-prioritisation occurs after data collection, usually without participants' knowledge or consent. However, our

results indicate that self-identification is strongly influenced by context, and may be constrained by factors such as racial stereotypes and institutional racism. Furthermore, asking for participants' main ethnic group may present a "forced choice dilemma", convey that multi-ethnic identity is unvalued, and lead to lower emotional wellbeing (Sanchez, 2010; Townsend et al., 2009). It is concerning that ethnic groups which have a greater tendency to be multi-ethnic (e.g., Māori and Pacific; Statistics New Zealand, 2020a) are disproportionately exposed to this dilemma and its negative consequences. This concern is particularly relevant to Māori because self-prioritisation is a colonial concept that is inconsistent with Māori concepts of ancestry (Jackson, 2003; Khawaja et al., 2000).

In research scenarios where potential strength of self-affiliation may be important to elicit, we propose that researchers revise the self-prioritisation question from "Which is your main ethnic group?" (which implies participants have one), to "Please indicate your main ethnic group(s) if you have one" (which does not make assumptions). Researchers also need to carefully consider how to classify multi-ethnic participants who do not select a main ethnic group. Based on the observation that most studies do not include this group as an explicit category, it can be inferred that researchers typically exclude these participants from their analytic sample. However, exclusion is problematic because the current study shows that, when given a number of response options, there is a significant proportion of multi-ethnic individuals who opt not to select a main ethnic group. In addition, present findings also suggest that non-selection is systematically associated with certain ethnic combinations (e.g., Asian/European) and lower socioeconomic deprivation. Therefore, exclusion of this group may lead to biased results.

## 5.6 Conclusion

In conclusion, this study clearly demonstrates that ethnic classification is a complex and subjective process with important implications for research and equity, especially given the steadily increasing global prevalence of multi-ethnic identifications. Therefore, it is imperative that quantitative researchers working with ethnic data recognise their position of power and critically consider the potential ramifications of their decisions regarding ethnic classification. Moreover, the additional limitations of self-prioritisation elucidated from this study indicate that, as the multi-ethnic population continues to grow, it will be important for researchers to review the use of this method, and progress ethnic classification in a way that best suits the evolving ethno-social context.

CHAPTER 6

# Study 3: Outcome Matters

The previous two empirical studies in this thesis demonstrate the importance of researchers' decisions on ethnic classification method in multi-ethnic contexts, because these decisions influence both outputted ethnic group sizes (Study 1), and the demographic composition of ethnic groups (Study 2). However, little is known about how ethnic classification method affects substantive outcomes in samples with higher multi-ethnic identification rates (e.g., >30%). Thus, Study 3 (this chapter) uses adolescent mental health outcomes as a case study to investigate this in a sample with 32% multi-ethnic prevalence. I chose to focus on adolescent mental health outcomes for four reasons: (1) this thesis was initially motivated by questions about the "best" ethnic classification method to use in adolescent mental health research; (2) the prevalence of mental health concerns substantially increases in adolescence due to biological, cognitive, and social changes (Thapar et al., 2012); (3) the effects of ethnic classification method on mental health is underexplored compared to physical health outcomes; and (4) the Youth'12 dataset has a relatively high multi-ethnic identification rate (compared to adults), and the ethnicity data are self-identified (child ethnicity data are via mother proxy).

Three specific mental health outcomes are examined in this study: (1) overall psychosocial difficulties (a numeric variable), (2) deliberate self-harm (binary variable with higher prevalence), and (3) suicide attempt (binary variable with lower prevalence). Examination of three outcomes will allow inferences on how ethnic classification effects are similar or different across outcomes in the same domain. All four common and accessible ethnic classification methods will be compared: total response (original and modified; see

description in Section 6.1.2.2), sole/combination grouping, administrative-prioritisation, and self-prioritisation. This study will inform researchers and policymakers of the implications that ethnic classification method has on conclusions regarding the needs of each ethnic group, the extent of inequities between ethnic groups, and consequently, decisions around intervention and resourcing.

As with the previous two chapters, the journal manuscript for this study is presented in its entirety below, with minor edits for consistency. The manuscript is published in the *Journal of Youth and Adolescence*, and the recommended citation is:

Yao, E. S., Bullen, P., Meissel, K., Tiatia, J., Fleming, T., & Clark, T. C. (2022). Effects of ethnic classification on substantive findings in adolescent mental health outcomes. *Journal of Youth and Adolescence*, *51*, 1581–1596. https://doi.org/10.1007/s10964-022-01612-6

## 6.1 Introduction

While most adolescents are healthy, mental health disorders affect around one in five young people and are one of the leading causes of morbidity and disability in adolescence (Clayborne et al., 2019; Thapar et al., 2012). Concerningly, epidemiological studies show that mental health challenges and associated indicators of distress, such as deliberate self-harm (i.e., non-suicidal self-injury) and suicide attempts, tend to disproportionately affect youth from Indigenous and ethnic minority backgrounds (Adkins et al., 2009; Anderson & Mayes, 2010; Benner et al., 2018; Crengle et al., 2012; Fleming, Tiatia-Seath, et al., 2020). However, classifying ethnicity for research in multi-ethnic contexts is complex. Previous studies, primarily conducted with adult physical health outcomes, suggest that researchers' choice of ethnic classification method can result in different substantive findings, and hence alter the implications and conclusions drawn (e.g., Boven et al., 2020; Callister & Blakely, 2004;

Lachowsky et al., 2020; Mays et al., 2003; Ministry of Health, 2008). However, little is known about how the choice of ethnic classification method affects results in the field of adolescent mental health. Therefore, the current study utilises a nationally representative adolescent sample from the multi-ethnic country of Aotearoa New Zealand to examine how different ways of classifying self-identified ethnicity affect substantive findings in the analysis of three mental health outcomes: overall psychosocial difficulties, deliberate self-harm, and suicide attempt. This will inform researchers in multi-ethnic contexts of the implications that their decisions regarding ethnic classification can have on the monitoring and addressing of ethnic inequities.

### 6.1.1 Ethnic Differences in Adolescent Mental Health Outcomes

Ethnicity—defined as socially-constructed groups with shared ancestry, history, traditions, culture, values, and beliefs (Morning, 2008)[24]—is a variable that has often been found to be associated with adolescent mental health outcomes, even after accounting for demographic characteristics such as socioeconomic deprivation. For example, a comprehensive review into internalising disorders in the United States indicates that racial/ethnic minority youth, particularly Hispanic youth, tend to have significantly poorer mental health outcomes when compared to non-Hispanic Whites (Anderson & Mayes, 2010). In Aotearoa New Zealand, a nationally representative survey series of secondary school students show that, in comparison to New Zealand European youth, Māori (Indigenous Peoples) and Pacific youth (those originating from neighbouring Pacific nations, e.g., Sāmoa, Cook Islands, and Tonga) tend to have significantly poorer mental health outcomes, including higher levels of mental health distress (Fleming, Tiatia-Seath, et al., 2020), self-harm (Fortune et al., 2010), and suicide attempts (Clark et al., 2018). The literature highlights racism at both

[24] The social construct of ethnicity is typically termed "race" in the United States, and will be referred to as such when U.S. research is referenced.

the individual level (e.g., prejudice and discrimination based on race/ethnicity) and institutional level (e.g., unequal distribution of opportunities and resources) as major contributing factors (Benner et al., 2018; Crengle et al., 2012; R. Harris et al., 2012; D. R. Williams, 2018). These mental health inequities are concerning from both developmental and equity perspectives alike.

Consequently, ethnicity is a variable that is crucial for monitoring, understanding, and addressing adolescent mental health concerns. First, ethnicity data are needed to monitor whether ethnic inequities are improving or worsening over time; for example, via trends in prevalence rates (e.g., Fleming, Tiatia-Seath, et al., 2020; Mojtabai et al., 2016). Second, ethnicity data are needed to understand the factors underlying ethnic disparities; for example, by testing models of risk and protective factors associated with each ethnic group (e.g., Adkins et al., 2009; Crengle et al., 2012). Third, ethnicity information is needed to target and evaluate interventions and policies aimed at decreasing ethnic inequities; for instance, by assessing whether an intervention is equally effective for each targeted ethnic group, or if it is underserving a particular group (e.g., Aspinall, 2018a; Mays et al., 2003).

Regression analysis is a statistical technique commonly used for ethnicity-based mental health analyses. A basic regression model usually specifies a mental health outcome as the dependent variable, and ethnicity and other relevant demographic characteristics as independent variables (Clark et al., 2018; Kessler et al., 2012; Mojtabai et al., 2016). These characteristics include: age, as there is typically a post-pubertal rise in mental health challenges due to biological, cognitive, and social changes (Thapar et al., 2012); sex, as adolescent females are around two times more likely to report mental health distress (Thapar et al., 2012); and the contextual characteristics of socioeconomic deprivation and urbanicity, as lower income and/or urban areas have been associated with poorer mental health outcomes (Fleming, Tiatia-Seath, et al., 2020; Kessler et al., 2012). More complex models may incorporate additional

independent variables, such as experiences of racism and discrimination, to seek to understand reasons underlying ethnic differences (Crengle et al., 2012).

However, many of these variables are social constructs (e.g., ethnicity, socioeconomic deprivation, urbanicity), so the way they are operationalised can influence results (Gillborn et al., 2018). This is particularly pertinent to ethnicity, as due to migration, interethnic unions, and changing patterns of self-identification, an increasing proportion of young people identify with more than one ethnic group (Aspinall, 2018b; Rocha & Aspinall, 2020). This raises the methodological issue of how to categorise multiple ethnic identifications for statistical analysis, and the question of how researchers' choice of ethnic classification method affects substantive findings in adolescent mental health. For example, it is not uncommon to see a statement like "Māori students [6.5%] were more likely than European students [2.7%] . . . to have attempted suicide in the past 12 months (*OR* = 1.88, 95% CI [1.35–2.60])" (Clark et al., 2018, p. 5). However, many studies do not provide clear information on how ethnicity has been operationalised, so it remains relatively unknown how different ethnic classification methods, such as the ones described in the subsequent section, change the reported "absolute" prevalence rate for an ethnic group (e.g., 6.5% for Māori), and the reported magnitude of relative difference between ethnic groups (e.g., odds ratio [*OR*] of 1.88 between Māori and European). Substantive variation in reported prevalence rates and magnitude of difference can impact conclusions on the level of mental health challenges experienced by each ethnic group, the extent of disparities between ethnic groups, and ultimately, how ethnic inequities are addressed via intervention and policy.

### 6.1.2 Classifying Multiple Ethnic Identifications

Ethnic classification methods for multiple ethnic identifications can be divided into two broad categories: *mutually exclusive methods*, where multi-ethnic participants are allocated to a single ethnic group; and *non-mutually exclusive methods*, where multi-ethnic participants are

allocated to two or more overlapping ethnic groups. Mutually exclusive methods, particularly ones that prioritise multiple ethnic identifications into broad ethnic groupings, tend to be more popular among applied researchers predominantly because they are easier to incorporate into statistical analysis (Callister et al., 2007; Mays et al., 2003; Yao et al., 2021). However, there are concerns associated with prioritisation methods because they suppress participants' multi-ethnic affiliation. Conversely, non-mutually exclusive methods preserve participants' multi-ethnic affiliation, but are more difficult to implement in statistical analysis.

As the specificities of ethnic classification methods are influenced by a country's ethno-cultural and socio-political context (Morning, 2008; Rocha & Aspinall, 2020), these methods will be discussed with reference to the broad ethnic groupings in the current study's context of Aotearoa New Zealand: European (e.g., New Zealand European [New Zealanders with European descent], White American, White British); Māori (Indigenous Peoples); Pacific Peoples (e.g., Sāmoan, Cook Islands Māori, Tongan); Asian (e.g., Chinese, Indian, Filipino); Middle Eastern, Latin American, and African (MELAA); and Other (residual category; Statistics New Zealand, 2005). Note that, except for Māori, these groupings are utilised for statistical output only, and contain considerable intragroup heterogeneity. The term "multi-ethnic" as used in this paper refers to individuals who self-identify with more than one of these broad ethnic groupings, although it should be noted that multiple identifications can also occur within each grouping (e.g., Sāmoan/Tongan). Aotearoa New Zealand is a valuable context for research on ethnic classification given its comparatively high ethnic diversity and rate of multi-ethnic identification, particularly among younger age groups (Statistics New Zealand, 2020a). In addition, the country has devoted considerable attention to ethnic classification in official statistics (Statistics New Zealand, 2004), and is obliged through its founding document, *Te*

*Tiriti o Waitangi* (*The Treaty of Waitangi*)[25] to ensure equity and address disparities between Māori and non-Māori (Cormack & Robson, 2010).

#### *6.1.2.1 Mutually Exclusive Methods*

Mutually exclusive ethnic classification methods include *sole/combination grouping*, *administrative-prioritisation*, and *self-prioritisation* (in the United States, these are respectively referred to as "multiracial combinations", "deterministic whole assignment", and "best race"; Office of Management and Budget [OMB], 2000). Each of these methods assign a participant to one ethnic category only. With sole/combination grouping, mono-ethnic participants are assigned to their sole ethnic group (e.g., sole European, sole Māori, sole Pacific, etc.), and multi-ethnic participants are assigned to their specific ethnic combination (e.g., Māori/European, Pacific/European, Māori/Pacific/European, etc.; Statistics New Zealand, 2005). With administrative-prioritisation, multi-ethnic participants are allocated to a single broad ethnic grouping according to a predetermined hierarchy. The standard hierarchy used in Aotearoa New Zealand is: Māori > Pacific > Asian > MELAA > Other > European (e.g., a Māori/European participant would be prioritised as Māori; Department of Statistics, 1993). With self-prioritisation, multi-ethnic participants are asked in a follow-up question to select one "main" ethnic group (Herman, 2004; Kukutai & Callister, 2009; Yao et al., 2021).

International literature tends to recommend sole/combination grouping because it retains participants' multi-ethnic identifications and allows more nuanced analysis (Aspinall, 2018a; Charmaraman et al., 2014; Sanchez et al., 2020). Sole/combination grouping is also an officially recommended output method by Statistics New Zealand (2005), alongside total response grouping (see Section 6.1.2.2). In contrast, administrative-prioritisation and self-

[25] *Te Tiriti o Waitangi* is a constitutional document between Indigenous Māori and the British Crown, whereby Māori are guaranteed the rights to partnership, participation, and protection.

prioritisation suppress multi-ethnic identifications—the former in a way that does not account for participants' strength of self-affiliation, and the latter in presenting a "forced choice dilemma" to participants (Shih & Sanchez, 2005; Statistics New Zealand, 2004; Yao et al., 2021). However, prioritisation methods tend to be more commonly used than sole/combination grouping in applied research in Aotearoa New Zealand. In particular, administrative-prioritisation tends to be routinely used in the education and health sectors (Cormack & Robson, 2010; Yao et al., 2021). This is likely due to a combination of factors, including recognition of Māori rights under *Te Tiriti o Waitangi* (Māori are prioritised in the hierarchy), and avoiding having a large number of ethnic categories, some with small subgroup sizes (Callister et al., 2007; Cormack & Robson, 2010; Yao et al., 2021).

Regardless of the specific ethnic classification method, inclusion of mutually exclusive ethnic groups as an independent variable for statistical analyses is relatively straightforward. For example, in a general linear model (GLM; e.g., analysis of variance [ANOVA], linear regression, logistic regression, multilevel modelling), the process usually involves selecting a reference group (typically the dominant ethnic group, e.g., European), dummy-coding the remaining groups, and then simultaneously entering the set of $k - 1$ ethnicity variables into the model ($k$ represents the total number of ethnic groups, the reference group is omitted from the model; Chatterjee & Simonoff, 2013; Cohen et al., 2003). Interpretation of resulting ethnicity coefficients relative to the reference group is reasonably intuitive, and because outcomes between the dominant ethnic group and minority ethnic groups are directly contrasted, effects of power and privilege versus marginalisation and disadvantage can be inferred.[26] Mutually

[26] Note some scholars argue that dummy-coding is problematic because the selected reference group is positioned as "normative", and have advocated for alternative coding methods (e.g., effect coding that compares each ethnic group with the overall sample; Mayhew & Simonoff, 2015; Ro & Bergom, 2020).

exclusive ethnic categories are also well-suited for, and simple to include in, statistical techniques such as chi-square tests of independence, multigroup confirmatory factor analysis (MGCFA), and structural equation modelling (SEM).

#### *6.1.2.2 Non-Mutually Exclusive Methods*

Alternatively, multiple ethnic identifications can be outputted using non-mutually exclusive methods, where a multi-ethnic participant is counted in each of the broad ethnic groups they identify with. The officially recommended—and commonly used—non-mutually exclusive ethnic classification method in Aotearoa New Zealand is *total response grouping* (Statistics New Zealand, 2005; known as the "all-inclusive method" in the United States; OMB, 2000). Under total response grouping, ethnicity data are structured as a set of six separate overlapping binary indicators which respectively indicates identification with each broad ethnic grouping (European: *yes/no*; Māori: *yes/no*; Pacific: *yes/no*; Asian: *yes/no*; MELAA: *yes/no*; Other: *yes/no*).

Statistical analyses using total response grouping typically involve a series of regression models containing one binary indicator at a time. Traditionally, each total response ethnic group of interest (*x*) is compared against all the participants who did not identify with that ethnic group (non-*x*; e.g., one model for Māori vs. non-Māori, one model for Pacific vs. non-Pacific, etc.). This method, hereafter referred to as *original total response*, results in reference groups with high intragroup heterogeneity, making it generally unsuitable for examining the effects of disadvantage and discrimination associated with ethnicity. For example, the reference group (e.g., non-Māori) usually includes participants who are also likely to experience marginalisation (e.g., Pacific, Asian, MELAA), and thus can attenuate observed effects. Another limitation is that the reference group is inconsistent across models.

Therefore, some researchers adopt a *modified total response* approach, where the series of regression models use a consistent reference group. The reference group usually comprises

the most privileged ethnic group(s) in the context of investigation. Examples of reference groups used in research in Aotearoa New Zealand include sole European (i.e., those who only identify as European; Clark et al., 2018; Lachowsky et al., 2020), sole European/Other (i.e., those who do not identify as Māori, Pacific, Asian, or MELAA; this differs from sole European because it includes responses such as "New Zealander"; Baker et al., 2012; Hobbs et al., 2019), and non-Māori/non-Pacific (i.e., those who do not identify as Māori or Pacific; Davis et al., 2006; Taua'i et al., 2018). While this approach addresses the limitations of the original total response method, it introduces new issues. For instance, because participants who are not in the comparison or reference groups are excluded from analyses, sample sizes are inconsistent across models, and not all available data are utilised.

Another non-mutually exclusive ethnic classification method, primarily described in the U.S. literature, is *fractional assignment* (Chatterjee & Simonoff, 2013; OMB, 2000). This method allocates weightings to each ethnic group selected by a participant so that they sum to 1, usually by assuming that a participant equally identifies with each selected group (e.g., for a Māori/European participant, 0.5 weighting will be assigned to Māori, and 0.5 weighting will be assigned to European). We do not consider fractional assignment an appropriate method to implement in the Aotearoa New Zealand context because it evokes parallels with the derogatory historic framing of multi-ethnic Māori through "blood quantum" terms (e.g., "half-caste", "quarter-caste"), and because Māori notions of whakapapa (ancestry) regard a person as fully belonging to each of their ancestries regardless of "blood quantum" (Jackson, 2003; A. D. Williams et al., 2018).

### 6.1.3 Effects of Ethnic Classification on Outcomes

The effects that ethnic classification method have on the analysis of outcomes is a relatively understudied area, especially in adolescent mental health. In Aotearoa New Zealand, the limited existing research has typically utilised adult datasets (<10% multi-ethnic

prevalence) to compare the effects of total response grouping, administrative-prioritisation, and sole/combination grouping on physical health outcomes—for example, smoking status (Boven et al., 2020), mortality (Callister & Blakely, 2004), sexual health (Lachowsky et al., 2020), and health indicators such as heart disease and diabetes (Ministry of Health, 2008). A noteworthy exception is Hobbs et al.'s (2019) study of ethnic classification effects on infectious diseases in early childhood (31% multi-ethnic prevalence), which also examined children's "self-prioritised" ethnicity as reported by their mother. In general, despite differences in sample and outcome measure, these studies indicate that prevalence rates within ethnic groups as delineated by total response grouping, administrative-prioritisation, and self-prioritisation (via mother report) differ by up to 5% before demographic characteristics are adjusted for. A slightly larger difference was occasionally observed between sole ethnic groups and total response ethnic groups, particularly for Māori (6% in Boven et al.'s [2020] study; 10% in Hobbs et al.'s [2019] study). After adjusting for demographic characteristics such as socioeconomic deprivation, ethnic classification method largely did not alter interpretations of whether an ethnic group was significantly different relative to the European referent ($p < .05$), but did at times produce marked variations in the magnitude of effect (e.g., *OR*s differed by up to 0.32 in Lachowsky et al.'s [2020] study; relative risk differed by up to 11.80 in Hobbs et al.'s [2019] study). Overall, the differences observed are noteworthy considering these arose solely due to a change in ethnic classification method.

Existing studies that delineate analyses by sole/combination grouping allow additional insight into the outcomes of multi-ethnic participants relative to their mono-ethnic counterparts. These show that, in Aotearoa New Zealand, socioeconomic and physical health outcomes for combination ethnic groups generally lie between their constituent sole ethnic groups (Boven et al., 2020; Callister et al., 2007; Callister & Blakely, 2004; Hobbs et al., 2019; Lachowsky et al., 2020). U.S. research with adolescents and adults (both <10% multiracial prevalence) has

observed similar patterns in socioeconomic outcomes (Bratter, 2018; Udry et al., 2003), self-rated health (Bratter & Gorman, 2011; Tabb et al., 2019), and educational outcomes (Cheng & Lively, 2009; Shih & Sanchez, 2005; Udry et al., 2003). In contrast, with mental health outcomes, U.S. research shows that multiracial youth generally tend to have similar or poorer outcomes on measures such as depression (Campbell & Eggerling-Boeck, 2006; Cheng & Lively, 2009; Fisher et al., 2014; Udry et al., 2003) and suicidality (Campbell & Eggerling-Boeck, 2006; Udry et al., 2003), when compared to their monoracial component group with lower psychological wellbeing, even after adjusting for socioeconomic background. The dominant explanation cited in the literature—which attributes these results to heightened identity conflict arising from having multiple heritages—is based on deficit discourse stemming from the anti-miscegenation era in the United States (Shih & Sanchez, 2005). Some contemporary scholars have importantly underscored the need to shift from an individualistic focus to addressing the wider social context where monoracial categories is the "norm", as this norm contributes to stigma, microaggression, and discrimination towards multiracial individuals (Gaither, 2015; Sanchez et al., 2020; Skinner et al., 2020). In Aotearoa New Zealand, the mental health status of multi-ethnic adolescents relative to their constituent ethnic groups is largely unknown.

## 6.2 The Current Study

Given the importance of ethnicity in adolescent mental health research and the dearth of information about the effects of ethnic classification methods in multi-ethnic contexts, this study utilises a large adolescent dataset from Aotearoa New Zealand to investigate how different ethnic classification methods affect the substantive findings of three adolescent mental health outcomes: (1) overall psychosocial difficulties, (2) deliberate self-harm, and (3) suicide attempt. The study was guided by two research questions:

1. How does ethnic classification method affect the *absolute level* of adolescent mental health outcomes reported for *each* ethnic group?
2. How does ethnic classification method affect the *relative difference* in adolescent mental health outcomes reported *between* ethnic groups?

## 6.3 Method

### 6.3.1 Data Source

Secondary data from the Youth'12 survey were utilised for this study. Youth'12 is a nationally representative cross-sectional survey of the health and wellbeing of secondary school students in Aotearoa New Zealand (aged 12–18 years; see Clark et al., 2013). Participants were selected using a two-stage clustered sampling design: one-third of secondary schools in the country were randomly selected, and within each of these schools, 20% of students on the school roll were randomly selected and invited to participate. In smaller schools (<150 students), 30 students from each school were randomly selected to protect confidentiality. Sampling weights were used to adjust for unequal likelihood of selection. Of the 12,503 randomly selected students, 8,500 (68%) participated in the survey. Among those who did not participate, the most common reasons were absence from school (22%), refusal to take part (20%), and unavailability due to other school activities (11%; Clark et al., 2013). Reason for non-participation was unavailable for 37% of non-participating students. The online survey was administered in schools via computer tablets, and was available in both English and te reo Māori (the Māori language). There was also optional audio voice-over. Ethical approval for Youth'12 was obtained from the University of Auckland Human Participants Ethics Committee (reference 2011/206).

Participants with missing ethnicity data ($n$ = 36; 0.4%), and/or other missing demographic data listed in the Measures section ($n$ = 192; 2.3%), were omitted from the current

study, resulting in an analytic sample of 8,275 (2.6% excluded in total). The small proportion of cases omitted (<5%) is unlikely to lead to biased results (Graham, 2009). Table 6.1 shows the overall demographic characteristics of the analytic sample. Thirty-two percent of the sample identified with more than one broad ethnic group. Demographic characteristics delineated by ethnicity, as classified using each ethnic classification method, are available in Supplementary Table C.1 in Appendix C.

**Table 6.1**

*Sample Demographic Characteristics (N = 8,275)*

| Demographic characteristic | *n* | % |
|---|---|---|
| Sex | | |
| Male | 3,752 | 45 |
| Female | 4,523 | 55 |
| Age (years) | | |
| ≤13 | 1,785 | 22 |
| 14 | 1,843 | 22 |
| 15 | 1,718 | 21 |
| 16 | 1,542 | 19 |
| ≥17 | 1,387 | 17 |
| Number of broad ethnic groups | | |
| 1 | 5,630 | 68 |
| 2 | 2,212 | 27 |
| ≥3 | 433 | 5 |
| Urbanicity | | |
| Main urban | 6,158 | 74 |
| Minor urban | 916 | 11 |
| Rural | 1,201 | 15 |
| New Zealand Deprivation Index (NZDep) | | |
| 1–2 (lowest deprivation) | 1,681 | 20 |
| 3–4 | 1,572 | 19 |
| 5–6 | 1,561 | 19 |
| 7–8 | 1,507 | 18 |
| 9–10 (highest deprivation) | 1,954 | 24 |

### 6.3.2 Measures

#### *6.3.2.1 Mental Health Outcomes*

Three measures of mental health were examined as dependent variables. *Overall psychosocial difficulties* was measured using total difficulties score in the Strength and Difficulties Questionnaire (SDQ; R. Goodman et al., 1998). Total difficulties score is the sum of the items in the SDQ's four difficulties subscales (emotional symptoms, peer problems, hyperactivity-inattention, and conduct problems), and has been found to be predictive of clinician-rated mental health diagnoses (A. Goodman & Goodman, 2009). Each subscale consisted of five items, and each item was rated on a 3-point Likert-type scale (0 = *not true*, 1 = *somewhat true*, 2 = *certainly true*). Negatively-worded items were reverse-scored. This produced a total difficulties score ranging from 0 to 40, with higher scores reflecting greater difficulties. The total score was then standardised to have a sample mean of 0 and standard deviation of 1. There were no partially missing responses, as the Youth'12 survey required participants to answer all the SDQ items (or skip the entire SDQ section).

*Deliberate self-harm* and *suicide attempt* in the past 12 months were dichotomous variables (0 = *no*, 1 = *yes*) respectively derived from responses to the following questions: "During the last 12 months, have you deliberately hurt yourself or done anything you knew might have harmed you (but not kill you)?", and "During the last 12 months, have you tried to kill yourself (attempted suicide)?". For the self-harm question, five response options were provided: *not at all* was coded as 0; and *yes—once*, *yes—two times*, *yes—3 to 5 times*, and *more than 5 times* were all coded as 1. For the suicide attempt question, four response options were provided: *not at all* and *not in the last 12 months* were coded as 0; and *once or twice* and *three or more times* were coded as 1.

#### *6.3.2.2 Ethnicity*

The focal independent variable of ethnicity, classified in five different ways, was based on self-report on two questions. The first question of "Which ethnic group do you belong to? (you may choose as many as you need)" had a check-all-that-apply format with 24 response options (e.g., New Zealand European, English, Australian, Māori, Sāmoan, Cook Island Māori, Filipino, Chinese, Indian, Middle Eastern, Latin American, African, etc.). Responses were aggregated into five broad total response ethnic groupings (European, Māori, Pacific, Asian, and Other),[27] then outputted in four ways: (1) *sole/combination grouping*, with twelve mutually exclusive dummy-coded categories (sole European [reference], sole Māori, sole Pacific, sole Asian, sole Other, Māori/European, Pacific/European, Asian/European, Māori/Pacific, Māori/Pacific/European, two groups not elsewhere included [NEI; e.g., Māori/Asian, Pacific/Asian], and three or more groups NEI [e.g., Māori/Pacific/Asian]); (2) *original total response grouping*, with five non-mutually exclusive binary indicators that compared each total response ethnic group *x* (European, Māori, Pacific, Asian, and Other) to the reference category of non-*x* (non-European, non-Māori, non-Pacific, non-Asian, and non-Other, respectively); (3) *modified total response grouping*, with four non-mutually exclusive binary indicators that compared each applicable total response ethnic group *x* (Māori, Pacific, Asian, and Other) to the identical reference category of sole European (note total response European was not applicable as a comparison group for this method because the reference group was sole European); and (4) *administrative-prioritisation*, with five mutually exclusive dummy-coded categories, where multiple ethnic responses were prioritised according to the following hierarchy: Māori > Pacific > Asian > Other > European (reference). The final classification method of (5) *self-prioritisation,* was based on the single-selection question of "Which is your

[27] MELAA was included in the "Other" category due to small subgroup size ($n$ = 248, 3%).

main ethnic group (the one you identify with most)?". The same list of 24 response options was used, with an additional "I can't choose only one ethnic group" option. Responses were aggregated into six mutually exclusively categories (European [reference], Māori, Pacific, Asian, Other, and "can't choose"), and then dummy-coded.

#### *6.3.2.3 Other Demographic Characteristics*

Other independent variables examined were the individual characteristics of *sex* (male [reference], female) and *age* in years (≤13 [reference], 14, 15, 16, ≥17), both self-reported; and the contextual characteristics of *urbanicity* and *neighbourhood socioeconomic deprivation*, both derived from participants' residential meshblock (a small geographic unit of approximately 60–120 residents based on participants' address). Urbanicity was classified as main urban (large urban areas with population ≥ 30,000 [reference]), non-main urban (small-medium urban areas with population between 1,000 and 29,999), and rural (population < 1000). Socioeconomic deprivation was measured using the New Zealand Deprivation Index (NZDep; Atkinson et al., 2014), a meshblock-based index constructed using 2013 census data such as income, employment, educational qualification, and home ownership. NZDep was grouped into quintiles, from the least deprived 20% of meshblocks (NZDep 1–2 [reference]) to the most deprived 20% of meshblocks (NZDep 9–10). Each of these variables were dummy-coded.

### 6.3.3 Data Analysis

Data were analysed using R version 4.0.2 (R Core Team, 2020). The relationship that the independent variables had with mental health outcomes, with a focus on ethnicity as classified in five different ways, was examined using multiple linear regression for total difficulties score, and binary logistic regression for self-harm and suicide attempt. Twelve separate regression models were specified for each outcome, with each model differing in the way ethnicity was classified (because membership in total response ethnic groups is not

mutually exclusive, separate models were specified for each of the five original total response indicators and each of the four modified total response indicators, resulting in nine models; in addition, one model was specified for each of the mutually exclusive ethnic classification methods: sole/combination grouping, administrative-prioritisation, and self-prioritisation). For each regression model, complete case analysis of the dependent variable was used, resulting in slightly different analytic sample sizes ($N$ = 7,990 [97%] for total difficulties score, 8,170 [99%] for self-harm, and 8,119 [98%] for suicide attempt). The small proportion of missingness (<5%) meant that complete case analysis was unlikely to lead to biased results (Graham, 2009). Note the modified total response models had smaller analytic sample sizes, as this classification method omits participants who are not in the comparison or reference groups.

From the resulting regression models, the "effects" package in R (Fox & Weisberg, 2019) was used to calculate adjusted mean estimates for total difficulties score, and adjusted prevalence estimates for self-harm and suicide attempt, for each ethnic group as classified by each ethnic classification method. The other independent variables were held constant at the sample average for these calculations. Note the residual ethnic category of "Other" was included in the regressions, but omitted from the figures in the results section, because it is a highly heterogenous group, making it difficult to draw inferences from.

## 6.4 Results

### 6.4.1 Descriptive Statistics

The overall sample mean for total difficulties score was a raw score of 11.37, 95% confidence interval (CI) [11.24, 11.49]. This was converted to a standardised score ($M$ = 0, $SD$ = 1) so that differences in scores are represented in standard deviation units (SDU; equivalent to $z$-scores and Cohen's [1988] $d$), and hence easier to interpret. The overall sample prevalence for deliberate self-harm in the past 12 months was 23.9%, 95% CI [23.0%, 24.8%]; and the

overall sample prevalence for suicide attempt in the past 12 months was 4.5%, 95% CI [4.0%, 5.0%]. Table 6.2 presents the regression estimates for each outcome by sole/combination ethnicity while keeping sex, age, urbanicity, and NZDep constant at the sample average (unadjusted estimates can be found in the rightmost columns of Supplementary Table C.1 in Appendix C). Note some combination groups (e.g., Māori/Pacific) had relatively small sample sizes, resulting in wider CIs. The significance levels in Table 6.2 show that, when compared to sole European, mental health outcomes tended to be significantly poorer for sole Māori, as well as for each ethnic combination examined.

At a descriptive level, combination ethnic groups (except Māori/European) tended to have higher total difficulties score, self-harm, and suicide attempt prevalence than their constituent sole ethnic groups. However, as indicated by the degree of overlap between their

**Table 6.2**

*Adjusted[a] Mental Health Outcomes by Sole/Combination Ethnicity*

| | | Total difficulties score[b] | | Self-harm (%) | | Suicide attempt (%) | |
|---|---|---|---|---|---|---|---|
| Ethnic group(s) | *n* | Estimate | 95% CI | Estimate | 95% CI | Estimate | 95% CI |
| European | 3907 | -0.07‡ | [-0.11, -0.04] | 22.0‡ | [20.7, 23.4] | 2.4‡ | [2.0, 3.0] |
| Māori | 288 | 0.07* | [-0.04, 0.19] | 22.7* | [18.1, 28.0] | 4.4* | [2.6, 7.2] |
| Pacific | 538 | -0.03 | [-0.12, 0.06] | 21.9 | [18.5, 25.8] | 5.7*** | [4.0, 7.9] |
| Asian | 738 | -0.08 | [-0.16, -0.01] | 17.0 | [14.4, 19.9] | 2.7 | [1.7, 4.1] |
| Other | 159 | 0.03 | [-0.12, 0.19] | 19.5 | [14.0, 26.5] | 3.6 | [1.6, 7.9] |
| Māori/European | 967 | 0.08*** | [0.02, 0.15] | 26.7*** | [24.0, 29.6] | 4.4** | [3.3, 5.8] |
| Pacific/European | 384 | 0.07* | [-0.03, 0.17] | 24.6* | [20.5, 29.3] | 6.2*** | [4.2, 8.9] |
| Asian/European | 224 | 0.06 | [-0.07, 0.19] | 23.3 | [18.1, 29.5] | 5.5** | [3.1, 9.5] |
| Māori/Pacific | 78 | 0.25** | [0.03, 0.47] | 23.7** | [15.7, 34.2] | 7.6** | [3.7, 14.8] |
| Māori/Pacific/European | 125 | 0.27*** | [0.09, 0.45] | 35.1*** | [27.0, 44.2] | 7.8*** | [4.3, 13.7] |
| 2 groups NEI | 559 | 0.06** | [-0.02, 0.15] | 27.8** | [24.2, 31.7] | 5.7*** | [4.1, 7.9] |
| ≥3 groups NEI | 309 | 0.33*** | [0.21, 0.44] | 34.8*** | [29.6, 40.5] | 7.9*** | [5.4, 11.4] |

*Note.* Significance levels denote statistically significant difference from the reference group based on multiple linear regression for total difficulties score, and binary logistic regression for self-harm and suicide attempt. CI = confidence interval; NEI = not elsewhere included.

‡Reference group. *$p < .05$. **$p < .01$. ***$p < .001$.

[a]Adjusted for sex, age, urbanicity, and NZDep. Estimates calculated at the weighted averages of these variables.

[b]Total difficulties score was standardised ($M = 0$, $SD = 1$).

95% CIs (Cumming & Finch, 2005), most of these differences were not statistically significant. In contrast, combination ethnic groups' demographic characteristics (e.g., NZDep) tended to lie between their constituent sole ethnic groups (see Supplementary Table C.1 in Appendix C).

### 6.4.2 Ethnic Classification Effects on Adjusted Estimates

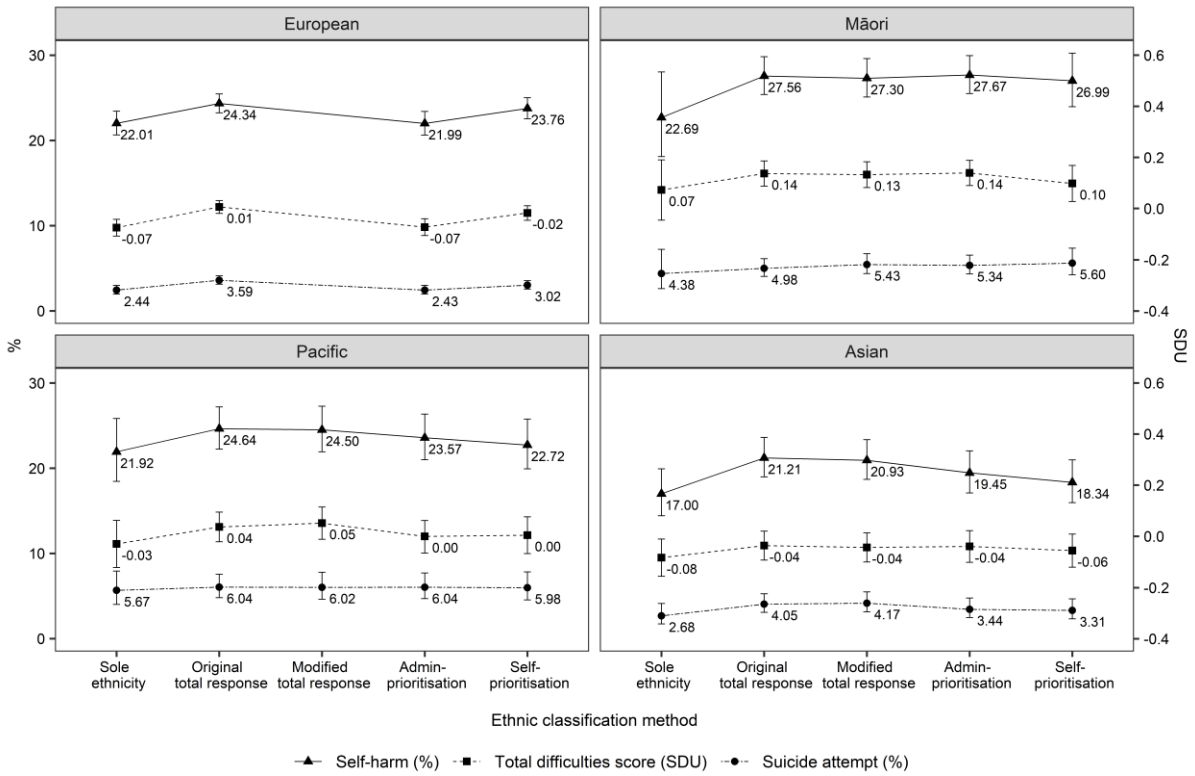
Figure 6.1 shows how adjusted estimates for mental health outcomes within each broad ethnic grouping fluctuated by ethnic classification method, as calculated using regression analyses. Due to space, for the sole/combination grouping method, only sole ethnic groups are shown for comparison (outcomes for combination groups can be found in Table 6.2). In general, sole ethnic groups had the lowest adjusted estimates (except for administratively-prioritised European, which was equivalent to sole European because European is in last position on the prioritisation hierarchy). Conversely, the two total response methods (original and modified) tended to have the highest adjusted estimates,[28] whereas the two prioritisation methods (administrative and self) tended to have estimates that were slightly lower than the total response methods. The estimates for administratively-prioritised Māori were an exception—these were similar to total response Māori due to its first position on the prioritisation hierarchy. Figure 6.1 also shows 95% CIs for each point estimate to indicate the range where there is relative certainty the true population value will lie. Note these CIs should not be used to infer statistical difference between ethnic classification methods within an ethnic group, as they do not represent independent groups (i.e., the data are dependent; Cumming & Finch, 2005).

Effect sizes were used to quantify the differences in adjusted estimates within each ethnic group. Cohen's (1988) *d*, which indicates the standardised difference between two

[28] Although original total response and modified total response included the same group of participants (e.g., total response Māori), their adjusted estimates varied slightly because a different reference group was used (e.g., non-Māori and sole European, respectively).

**Figure 6.1**

*Adjusted[a] Mental Health Outcomes Within Ethnic Groups by Ethnic Classification Method*




*Note.* Error bars show 95% confidence intervals. Total difficulties score was standardised ($M = 0$, $SD = 1$), so is represented in standard deviation units (SDU).

[a]Adjusted for sex, age, urbanicity, and NZDep. Estimates calculated at the weighted averages of these variables.

means, was used for total difficulties score. For prevalence differences in deliberate self-harm and suicide attempt, Cohen's *h* (an effect size for proportions analogous to Cohen's *d* for means) was used. These effect size measures are appropriate for both independent and dependent groups (Cohen, 1988). For reference, Cohen's rule of thumb for interpreting *d* and *h* are: 0.20 = small effect, 0.50 = medium effect, and 0.80 = large effect. It is important to note that this interpretation is usually applied to differences between groups (e.g., Māori vs. European), rather than differences due to methodological changes alone (e.g., sole Māori vs. total response Māori), so the benchmarks are overly conservative for our purposes.

In this study, the largest within-group differences resulting from a change in ethnic classification method tended to have an effect size between 0.05 and 0.10. For example,

adjusted suicide attempt prevalence for sole European was 2.4%, compared to 3.6% for total response European. This translates to a percentage increase of 50% and had an effect size of $h = 0.07$. For Māori, the largest difference in adjusted suicide attempt prevalence was between sole Māori (4.4%) and self-prioritised Māori (5.6%), equating to a 27% increase and an effect size of $h = 0.06$. The largest within-group difference in effect sizes was observed in Māori self-harm rates—the rate was 22.7% for sole Māori (note relatively large CI) and 27.7% for administratively-prioritised Māori (relatively smaller CI), reflecting a 22% increase and an effect size of $h = 0.12$.

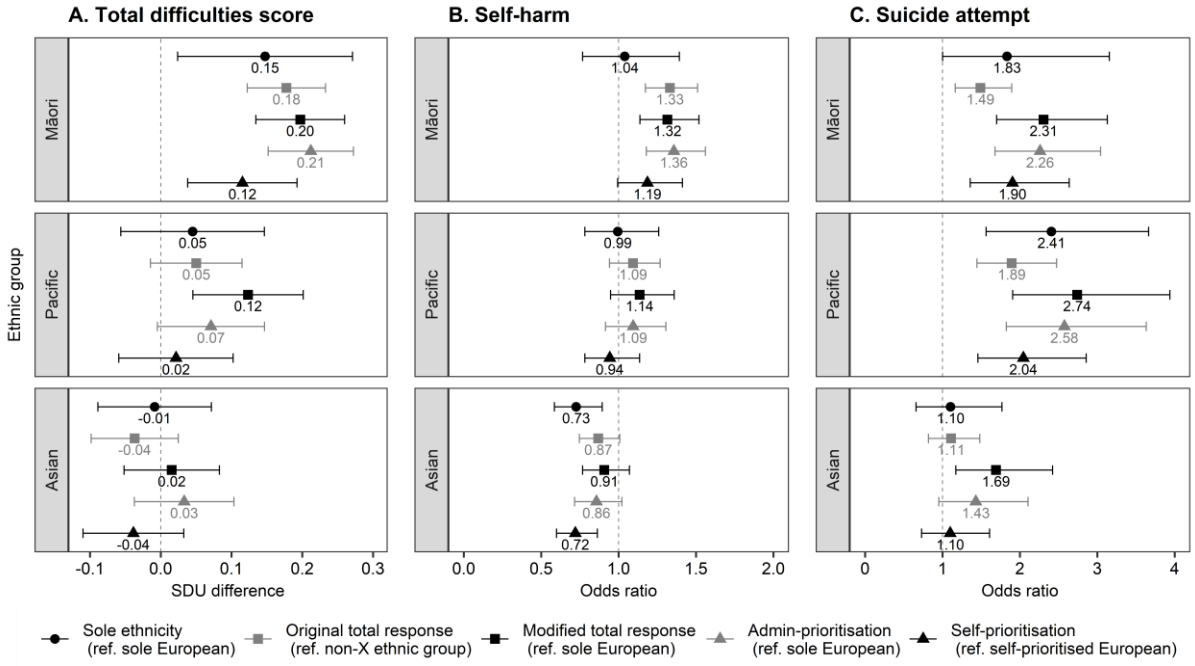
### 6.4.3 Ethnic Classification Effects on Subgroup Differences

Next, the effects of ethnic classification method on between-group differences in mental health outcomes, as estimated using regression analyses, were investigated (again, for sole/combination ethnicity, only sole ethnic groups were used for comparison). Figure 6.2 shows the partial regression coefficients and 95% CIs for ethnicity, after adjusting for sex, age, urbanicity, and NZDep (full results, which include parameter estimates for all demographic characteristics, can be found in Supplementary Tables C.2 to C.7 in Appendix C). First, we examined whether a change in ethnic classification method altered the interpretation of whether an ethnic group's mental health outcome was significantly different to its reference group ($p < .05$, indicated by a 95% CI that does *not* cross the dotted null effect line; as above, CIs should not be compared across ethnic classification methods due to their dependency; Cumming & Finch, 2005). The reference group for each classification method is shown on the legend.

The CIs in Figure 6.2 show that ethnic classification method did not alter significance interpretations when ethnicity effects were relatively large. For example, suicide attempt prevalence for Māori and Pacific was significantly higher than their respective reference group regardless of the ethnic classification method used (none of these CIs crossed the null effect

**Figure 6.2**

*Partial Regression Coefficients[a] for Ethnicity by Ethnic Classification Method and Mental Health Outcome*




*Note.* Error bars show 95% confidence intervals (note different *x*-axis scale for each outcome). Vertical dotted line indicates the null effect line. European panel is not applicable as it was the reference group in most cases. Total difficulties score was standardised ($M = 0$, $SD = 1$), so group differences are represented in standard deviation units (SDU).

[a]Controlling for sex, age, urbanicity, and NZDep.

line). However, when ethnicity effects were smaller, significance interpretations were sometimes inconsistent between ethnic classification methods (some CIs crossed the null effect line, others did not). For example, total difficulties score for Pacific was significantly different from its referent when modified total response was used, but no significant differences were observed under the remaining four ethnic classification methods. Similarly, self-harm for Māori was significantly different from its referent according to total response (both original and modified) and administrative-prioritisation, but not significantly different according to sole ethnicity and self-prioritisation. Note sole ethnicity typically had wider CIs than the other classification methods due to smaller subgroup sizes.

We also examined the extent that the magnitude of effects for ethnicity differed by ethnic classification method (indicated by the distance between points within each sub-panel in Figure 6.2). For all three mental health outcomes, modified total response and administrative-prioritisation tended to produce the largest ethnicity effects for each ethnic group, whereas sole ethnicity and self-prioritisation tended to produce smaller effects. However, the impact of original total response depended on the outcome—it produced relatively larger effects for self-harm, and relatively smaller effects for total difficulties score and suicide attempt. Differences in the magnitude of effect by ethnic classification method were most substantial for suicide attempt. For example, for both Māori and Pacific, the smallest adjusted *OR* (original total response; 1.49 and 1.89, respectively) and largest *OR* (modified total response; 2.31 and 2.74, respectively) differed by more than 0.80, equivalent to a Cohen's (1988) *d* effect size of up to 0.25.[29] For total difficulties score and self-harm, the largest difference in magnitude within each ethnic group was less substantial but still noteworthy (*d* ranged from 0.07 to 0.10, and 0.10 to 0.15, respectively).

The effect of ethnic classification method on the results of the other demographic characteristics was not a key focus of this study. However, a brief examination of the significance patterns at $p < .05$ between the different models for each outcome (see Supplementary Tables C.2 to C.7 in Appendix C) showed that females and those living in higher NZDep areas (i.e., poorer neighbourhoods) had significantly worse mental health outcomes regardless of the ethnic classification method used. Within each outcome, significant differences by age tended to be relatively consistent across ethnic classification methods. Outcomes were generally not significantly associated with urbanicity regardless of the ethnic

[29] *OR*s were converted to Cohen's (1988) *d* effect size using the formula of $ln(OR)/1.81$ (Chinn, 2000). This allows for comparison of ethnic classification effects on a common scale.

classification method used. The only exception was for self-harm, where the modified total response Māori model and the self-prioritised model indicated that those living in rural areas reported significantly lower self-harm rates than those living in urban areas.

## 6.5 Discussion

Ethnicity is an important variable for adolescent mental health research, but little is known about how different ways of classifying self-identified ethnicity impact the conclusions drawn. The current study investigated how common ethnic classification methods affect the substantive findings of adolescent mental health outcomes in a nationally representative sample in Aotearoa New Zealand with 32% multi-ethnic prevalence. Overall, the results indicate that the majority of adolescents did not have significant mental health concerns. However, around one in five reported deliberate self-harm, and one in twenty reported attempting suicide, in the year prior to the survey. Consistent with existing patterns in Aotearoa New Zealand, Māori and Pacific youth generally tended to have disproportionately inequitable mental health outcomes in comparison to European and Asian youth (Clark et al., 2018; Fleming, Tiatia-Seath, et al., 2020). However, the current study revealed important nuances by ethnic classification method and outcome measure, demonstrating the influence of methodological decisions on research conclusions in multi-ethnic contexts.

### 6.5.1 Ethnic Classification Effects Within Ethnic Groups

Within each broad ethnic grouping, mental health outcomes tended to fluctuate by the ethnic classification method used. The largest differences typically ranged between a Cohen's (1988) *d* and *h* effect size of 0.05 and 0.10 (the largest observed effect size was 0.12), which we consider substantial given this was solely due to a change in ethnic classification method within the same sample. Within each ethnic group, sole ethnicity generally produced the most positive mental health outcomes out of the five ethnic classification methods examined (e.g.,

lower self-harm prevalence), and total response methods (i.e., original and modified total response) generally produced the least positive outcomes (e.g., higher self-harm prevalence). Outcomes as analysed by the prioritisation methods (i.e., administrative-prioritisation and self-prioritisation) tended to lie in-between (note exceptions for administratively-prioritised Māori and European due to their first and last position on the prioritisation hierarchy, respectively). In other words, studies that use different ethnic classification methods can potentially report substantively different descriptive results. For example, using the same dataset, a study which examined adjusted self-harm rates for Māori using sole ethnicity would have reported a rate of 22.7%, whereas a study which used original total response would have reported a rate of 27.6%. Therefore, consistent with previous studies in Aotearoa New Zealand which investigated physical health outcomes among children (Hobbs et al., 2019) and adults (Boven et al., 2020; Ministry of Health, 2008), descriptive statistics of mental health outcomes were influenced by ethnic classification method. However, it is important to note that the exact effect of ethnic classification method depends on the subsample and outcome measure, regardless of whether the outcomes are within a particular domain (e.g., the current study showed that the magnitude of effect differed somewhat across ethnic groups and mental health outcomes), or across different domains (e.g., the current study suggests sole Māori had lower rates of self-harm than total response Māori, whereas Boven et al.'s [2020] study with adults showed that sole Māori had higher rates of tobacco smoking than total response Māori).

Nuance provided through sole/combination grouping suggests that, in this study, the pattern of results by ethnic classification method may be attributable to how multi-ethnic participants were classified. Specifically, total response grouping includes all participants who identified with the specified ethnic group (whether alone ["mono-ethnic"] or in combination ["multi-ethnic"]), the prioritisation methods may only classify some multi-ethnic participants to the ethnic group (e.g., under administrative-prioritisation, the Pacific grouping would

exclude Māori/Pacific participants, but include Pacific/European participants), and sole/combination grouping separates sole-ethnic and multi-ethnic participants into specific categories. The proportion of multi-ethnic participants included may lead to different reported outcomes by ethnic classification method because, while differences between multi-ethnic combinations and their constituent ethnic groups were typically not statistically significant, combination groups (with the exception of Māori/European) tended to have poorer mental health outcomes than their constituent groups at a descriptive level. These mental health disparities are similar to patterns observed among adolescents in the United States in regards to depression (Campbell & Eggerling-Boeck, 2006; Cheng & Lively, 2009; Fisher et al., 2014; Udry et al., 2003) and suicidality (Campbell & Eggerling-Boeck, 2006; Udry et al., 2003), and are likely driven by societal stigma (e.g., negative stereotypes) and discrimination (e.g., identity denial or questioning) towards multi-ethnic individuals (Sanchez et al., 2020; Skinner et al., 2020). As above, it is important to note that these patterns may differ across outcomes. For example, consistent with previous research with adults in Aotearoa New Zealand (Boven et al., 2020; Lachowsky et al., 2020) and the United States (Bratter, 2018; Udry et al., 2003), the current study found that, unlike mental health outcomes, the demographic characteristics (e.g., socioeconomic profile) of ethnic combination groups tended to lie between their component ethnic groups.

### 6.5.2 Ethnic Classification Effects Between Ethnic Groups

In addition to reporting descriptive statistics by ethnic group, ethnicity data are also frequently used to monitor and address inequities between groups (Fleming, Tiatia-Seath, et al., 2020; Mays et al., 2003; Mojtabai et al., 2016). When investigating ethnic inequities, interpretations are typically drawn by examining: (1) whether there is a significant difference in outcomes between an ethnic group and a referent (typically the dominant ethnic group), and (2) the magnitude of difference between the groups. Similar to Lachowsky et al.'s (2020) study

on the effects of ethnic classification on sexual health outcomes for homosexual men in Aotearoa New Zealand, our results on adolescent mental health show that conclusions on whether ethnic groups are significantly different can depend on the ethnic classification method and outcome measure. For example, in comparison to the reference group (i.e., non-*x* ethnic group for original total response, self-prioritised European for self-prioritisation, and sole European for all other methods), three of the five classification methods examined showed that Māori youth had significantly higher self-harm prevalence, and one of the five methods showed Pacific youth had significantly higher total difficulties score. In contrast, all five ethnic classification methods examined showed that Māori and Pacific adolescents had higher rates of suicide attempt and that Māori had higher total difficulties score. In addition, the prevalence of self-harm in Pacific youth was not significantly different from their peers regardless of classification method. Based on patterns of ethnic group differences for each outcome, it appears that ethnic classification method is more likely to influence significance interpretations when the magnitude of ethnic differences is small, but interpretations are relatively robust when the magnitude is large. It is also important to note that explanatory power can differ by ethnic classification method. For example, because Māori and Pacific adolescents tend to have higher rates of multi-ethnic identification, there were fewer participants in the respective sole ethnic groups, resulting in larger error margins and less statistical power to detect an effect if it exists (i.e., higher chance of Type II error).

In terms of magnitude of ethnic differences in adolescent mental health, modified total response and administrative-prioritisation tended to result in the largest effects, likely because these methods compare groups that include multi-ethnic participants (e.g., total response Māori, total response Pacific)—who at a descriptive level tended to have poorer mental health outcomes relative to their constituent ethnic groups—to the reference group of sole European. For example, for ethnic differences in suicide attempt for Māori and Pacific youth, changing

the reference group from sole European (i.e., modified total response) to non-Māori and non-Pacific (i.e., original total response), respectively, reduced the magnitude of between-group differences by an effect size (*d*) of up to 0.25. This indicates the influence of researchers' choice of reference group, and has important implications particularly for health and education research in Aotearoa New Zealand, where administrative-prioritisation is routinely used (Cormack & Robson, 2010; Yao et al., 2021). Although the referent of sole European may be appropriate for studies interested in the impact of ethnic marginalisation on outcomes (Lachowsky et al., 2020), depending on the outcome, it is possible that observed differences may be partially attributable to contextual factors that are detrimental to multi-ethnic youth, such as multi-ethnic stigma and discrimination (Sanchez et al., 2020; Skinner et al., 2020). These factors can act as confounds in analyses if not explicitly modelled.

Ethnic differences in mental health outcomes tended to be smaller when self-prioritised ethnicity was used. This may partly be due to the inclusion of multi-ethnic participants in both the comparison and reference groups, but because self-prioritisation can arguably be a crude proxy of strength of ethnic affiliation (Atatoa Carr et al., 2017; Kukutai & Callister, 2009; Yao et al., 2021), it is also possible that it reflects the protective effect of strong ethnic identity on mental health (Anderson & Mayes, 2010; Fisher et al., 2014; A. D. Williams et al., 2018). However, research suggests that self-prioritisation is influenced by contextual factors such as societal stereotypes (Herman, 2004; Yao et al., 2022). Therefore, if researchers are interested in the effects of ethnic identity, it would be more appropriate to measure this directly, for example, using the Multigroup Ethnic Identity Measure (MEIM; Phinney & Ong, 2007), for each of the ethnic groups identified by a participant.

In summary, researchers' conclusions about ethnic group differences, both in terms of statistical significance and magnitude of effect, can differ depending on the ethnic classification method. Moreover, the effect of method can differ by outcome measure, even if these measures

are all within the mental health domain. Finally, it is important to note that while ethnic classification methods had an influence on the interpretation of ethnic group differences, other demographic characteristics (e.g., sex, age, socioeconomic deprivation) were relatively robust to changes in ethnic classification method.

### 6.5.3 Limitations and Future Research

Some limitations of this study need to be noted. First, the study was conducted with a cross-sectional adolescent sample from 2012, as this survey included a question on self-prioritised ethnicity. Although it was a nationally representative sample at the time, the 2019 iteration of the survey shows that the proportion of participants who identified with an Asian ethnic group increased by over 10% from 2012 (Fleming, Peiris-John, et al., 2020), and that, consistent with international trends, there has been an overall rise in internalising mental health symptoms (Fleming, Tiatia-Seath, et al., 2020). This may have some impact on the generalisability of the study, although initial reports indicate that the rate of multi-ethnic identification, the proportions of total response ethnic groups other than Asian, as well as the relative mental health status of total response ethnic groups (including Asian), were similar across the two cohorts (Fleming, Peiris-John, et al., 2020; Fleming, Tiatia-Seath, et al., 2020).

Second, total difficulties from the SDQ was examined as a composite score rather than a latent factor with inherent measurement error. While this is a common way of using the well-validated SDQ (Achenbach et al., 2012), to our knowledge the measure has not been validated among adolescents in Aotearoa New Zealand. Future research could explore ethnic invariance of the SDQ in this age group and context, including whether conclusions differ by ethnic classification method, as this will provide valuable information for researchers who wish to work with ethnicity data under a latent framework.

Third, the focus of the present study was to investigate the effects that ethnic classification method has on the substantive findings of ethnic differences in adolescent mental

health outcomes, rather than to explore possible mechanisms underlying these differences. This risks the erroneous interpretation that observed disparities are due to inherent differences rather than factors such as systemic disadvantage experienced by Indigenous and ethnic minority groups. While it was beyond the scope of this study, we underscore the importance of research that directly examines causal factors (e.g., ethnic identity and racial discrimination; Crengle et al., 2012; R. Harris et al., 2012; A. D. Williams et al., 2018) associated with mental health outcomes for specific ethnic groups, in order to inform knowledge, intervention, and policy. It would be interesting for future research to examine whether these results differ by ethnic classification method.

### 6.5.4 Implications

Despite these limitations, the effects that ethnic classification method had on substantive outcomes in the current study have important implications for research and practice both in Aotearoa New Zealand and internationally, especially given the rapid growth of the multi-ethnic population worldwide (Aspinall, 2018b; Rocha & Aspinall, 2020). This study highlights the influential role that researchers, through their choice of ethnic classification method, have on knowledge construction in the increasingly diverse global context. There are three key research implications from this study. First, it is imperative that researchers critically select their ethnic classification method when undertaking research that includes ethnicity as a variable, and collect ethnicity data in a way that enables the method to be utilised. Second, researchers need to explicitly state the ethnic classification method chosen, both for transparency and to facilitate the study's replicability. Ideally, this should be accompanied by the rationale and possible implications the method may have on results. Third, it is important to educate research audiences—ranging from students and researchers, to policymakers and the media—on the complexities of ethnic classification, encourage them to engage more

critically with research, and understand that the conclusions drawn can depend on the methods used.

In terms of critically selecting an ethnic classification method, there are two major aspects researchers need to carefully consider. First, researchers should select the method most suitable to the research question and context. It is helpful here to think about who the comparison and reference groups should include. For example, in the Aotearoa New Zealand context, if the research purpose is to examine equity between Māori and European as guaranteed by *Te Tiriti o Waitangi* (*The Treaty of Waitangi*), administrative-prioritisation or modified total response would be more appropriate, because these methods compare everyone who identifies as Māori to those who solely identify as European. In contrast, if the purpose is to examine equity for Pacific Peoples, modified total response would be more appropriate than administrative-prioritisation, because administrative-prioritisation subsumes those who identify as both Pacific and Māori into the Māori category. If the purpose is to examine ethnicity with more nuance, sole/combination grouping may be appropriate, with the caveat that these results need to be interpreted with extreme care. In particular, sole/combination ethnicity tends to be highly fluid, and some categories will have small subgroup sizes, resulting in unequal explanatory power (Aspinall, 2018a; Callister et al., 2007; Cormack & Robson, 2010). Moreover, the relatively individualised results need to be interpreted in a way that does not place blame on individuals or promote deficit thinking, but rather, considers the wider sociohistorical context (e.g., stigma and discrimination towards certain groups).

Second, researchers conducting studies which include Indigenous Peoples should adhere to the four *CARE Principles for Indigenous Data Governance* (Research Data Alliance International Indigenous Data Sovereignty Interest Group, 2019) when selecting their ethnic classification method, such that they: (1) contribute to the *Collective Benefit* of Indigenous Peoples; (2) allow Indigenous Peoples the *Authority to Control* decisions made regarding

Indigenous data, including how ethnicity is classified; (3) fulfil their *Responsibility* in sharing how their research supports Indigenous Peoples; and (4) uphold *Ethics* so that research findings using the selected ethnic classification method both minimises harm, and maximises benefit, to Indigenous Peoples. The general principle of minimising harm and maximising benefit should likewise extend to other ethnic groups traditionally marginalised in research (e.g., Pacific and Asian).

While it was not the main focus of this study, the study also has important implications for practice and policy both in Aotearoa New Zealand and other ethnically diverse countries. In particular, the current study generally supports existing research that Māori and Pacific adolescents, and adolescents who identify with more than one ethnic group, tend to have higher mental health needs in comparison to sole European youth (Cheng & Lively, 2009; Clark et al., 2018; Fisher et al., 2014; Fleming, Tiatia-Seath, et al., 2020). The literature indicates these inequities are largely due to interpersonal and systemic racism (Benner et al., 2018; Crengle et al., 2012; R. Harris et al., 2012; D. R. Williams, 2018). Therefore, at the individual and community levels, it is important that parents, practitioners, schools, and communities support Indigenous, ethnic minority, and multi-ethnic youth during the crucial adolescent developmental phase to develop positive ethnic identity, and equip them with tools to buffer the negative impacts of racial stigma, stereotypes, and discrimination (Benner et al., 2018; Sanchez et al., 2020). Equally, continual attention is needed at the societal level to address the systemic disadvantage and marginalisation that contribute to ethnic inequities in mental health (Sanchez et al., 2020; A. D. Williams et al., 2018).

## 6.6 Conclusion

Ethnicity data are critical for monitoring and addressing ethnic inequities, but little is known about how different ethnic classification methods impact substantive findings,

particularly in adolescent mental health research. Using a nationally representative adolescent sample with over 30% multi-ethnic prevalence, the current study empirically demonstrates via three mental health outcomes that different ethnic classification methods can lead to different substantive results. Most notably, solely due to the ethnic classification method used, reported mental health outcomes within the same nominal ethnic group varied by an effect size (*d*) of up to 0.12, and the reported magnitude of difference between nominal ethnic groups varied by an effect size (*d*) of up to 0.25. Therefore, it is paramount that researchers exercise criticality and transparency when working with ethnicity data, because their decisions impact the conclusions drawn; which in turn influence intervention, policy, and practice; and ultimately, the health and wellbeing of young people.

CHAPTER 7

# General Discussion: Interpretation Matters

The motivation for the three studies in this thesis arose from the overarching aim to empirically investigate the implications that researchers' decisions on ethnic classification method have on quantitative analysis in the increasingly multi-ethnic global context. Ethnicity is a crucial variable for research and policy for equity and social justice, but is complex to operationalise for quantitative research given its fluidity, subjectivity, and multiplicity (Atatoa Carr et al., 2017; Balestra & Fleischer, 2018; Callister et al., 2007; Mays et al., 2003). Since the 2000 census round, there has been a shift in many countries towards collecting ethnicity data by asking respondents to select all the ethnic groups they identify with, both in official enumeration and in surveys (Balestra & Fleischer, 2018; Morning, 2008; United Nations, 2008). This poses the issue of how to output multiple ethnicity data for statistical analysis, especially given that a small number of mutually exclusive categories is typically preferred for nominal variables (Cohen et al., 2003; Field et al., 2012).

A number of possible ethnic classification methods have been outlined in the literature (Cormack & Robson, 2010; Herman, 2011; OMB, 2000; Statistics New Zealand, 2004), and can broadly be grouped into methods that retain multiple ethnicity data (e.g., total response and sole/combination grouping), and methods that reduce multiple ethnicity data (e.g., administrative-prioritisation and self-prioritisation). Each method has associated strengths and limitations, so there is no universally recognised "best" method (Atatoa Carr et al., 2017; Moubarac, 2013; OMB, 2000; Woo et al., 2011). Previous research, mainly using samples with relatively low multi-ethnic prevalence (e.g., <10%), indicates that ethnic classification method can impact quantitative results (Boven et al., 2020; Callister et al., 2007; Mays et al., 2003;

Rutkowski et al., 2017). However, little is known about the extent of this impact on more multi-ethnic samples. In addition, there appears to be a disconnect between the existing literature and mainstream research practice, because many researchers do not make their method of ethnic classification transparent, suggesting that they are either unaware of, or do not critically consider, its potential implications (Moubarac, 2013).

Therefore, this thesis utilised large-scale survey data of children, adolescents, and adults from Aotearoa New Zealand to investigate the following overarching research question: *how does ethnic classification method affect applied quantitative analysis in multi-ethnic contexts?* The effects of ethnic classification method were examined in three areas, corresponding to the three studies presented in Chapters 4 to 6, respectively: (1) outputted ethnic group size, (2) the demographic composition of outputted ethnic groups, and (3) substantive outcomes in mental health. I approached this research from a critical quantitative perspective, with the dual goals of addressing the gap in the literature, as well as addressing the disconnect between the literature and mainstream practice. The three studies are synthesised and discussed holistically in this general discussion chapter. The chapter will first summarise the main findings from each study, then discuss the four common ethnic classification methods in light of these findings and the wider literature. Thereafter, the limitations of this thesis are outlined, the overall implications for working with ethnicity data are presented, and the contributions of this thesis are discussed.

## 7.1 Summary of Main Findings

In Study 1 (Chapter 4), I examined how researchers' choice of ethnic classification method influences outputted ethnic group size in increasingly multi-ethnic cohorts—adults (16% multi-ethnic), adolescents (32% multi-ethnic), and children (38% multi-ethnic). Ethnic group size is important because it represents the participants who are included or excluded in

the ethnic group, constitutes the denominator in ethnic-specific analyses, and is associated with resource allocation decisions. Results showed that ethnic classification method matters towards outputted ethnic group sizes—as expected, administrative-prioritisation and self-prioritisation tended to understate ethnic group sizes compared to total response grouping. The effects of ethnic classification method were particularly pronounced in cohorts with higher multi-ethnic identification (i.e., children and adolescents), and in the European and Māori groupings (groups at either end of the administrative-prioritisation hierarchy). For example, in the child and adolescent samples, when compared to total response grouping, administrative-prioritisation and self-prioritisation reduced the European proportion by over 10% and 25%, respectively; and self-prioritisation halved the outputted Māori proportion (the proportion of administratively-prioritised Māori was identical to total response Māori due to the group's first position on the prioritisation hierarchy). The marked effect that ethnic classification method can have on ethnic group counts, particularly in samples with higher multi-ethnic prevalence, highlights the importance of researchers' ethnic classification decisions.

In addition, Study 1's examination of multi-ethnic participants' responses to the self-prioritisation question showed that while the majority in each age cohort selected a main ethnic group, around 20% could not or refused to do so, and there was over 60% discrepancy between self-prioritised ethnicity and administratively-prioritised ethnicity. This indicates that these two popular ethnic classification methods are not interchangeable, and raises the question of whether researchers' choice of prioritisation method results in differences in the demographic composition of outputted ethnic groups. Thus, Study 2 (Chapter 5) explored the demographic characteristics that are associated with discrepancies between administrative-prioritisation and self-prioritisation. Results showed that, in each cohort, discrepancy rates due to participants selecting a main ethnic group that differed from their administratively-prioritised ethnicity were systematically associated with the contextual characteristics of neighbourhood ethnic

composition and socioeconomic deprivation, but largely not associated with the individual characteristics of age, sex, and birthplace. In the context of a dual-ethnic combination comprising an Indigenous/ethnic minority group and the ethnic majority group (e.g., Māori/European), this means that, under self-prioritisation, these individuals are more likely to be classified to an Indigenous/ethnic minority group if they lived in neighbourhoods of higher ethnic minority density or higher socioeconomic deprivation, and vice versa under administrative-prioritisation. In other words, researchers' choice of ethnic classification method can bias samples and subsequent analyses. In addition, for children and adolescents, discrepancies due to participants (or their caregivers) not selecting a main ethnic group was more likely to occur in those living in less socioeconomically deprived neighbourhoods. In summary, the results from Study 2 suggest that context matters in multi-ethnic individuals' choice of a main ethnic group, as well as decisions to not select a main ethnic group.

In Study 3 (Chapter 6), I shifted focus from examining the effects of ethnic classification on sample size and demographic composition, to its effects on substantive outcomes in adolescent mental health. Results showed that, due solely to changes in ethnic classification method, reported outcomes within the same nominal ethnic group varied by an effect size (*d*) of up to 0.12. Sole ethnicity tended to produce the most positive mental health outcomes, followed by the prioritisation methods (administrative-prioritisation and self-prioritisation), then the total response methods (original total response and modified total response). For between-group differences, results showed that ethnic classification method can influence conclusions on whether there were significant disparities between two nominal ethnic groups, and also the magnitude of difference between them, up to an effect size (*d*) of 0.24. Modified total response and administrative-prioritisation tended to produce larger between-group differences compared to original total response and self-prioritisation. However, the outcome measure matters when investigating ethnic classification effects, as there were

variations in these general patterns even when the three outcomes examined were all within the domain of mental health. Results for other demographic characteristics (e.g., age, sex, and socioeconomic deprivation) remained relatively robust to changes in ethnic classification method.

Together, the three studies demonstrate the power of researchers, because the ethnic classification method they choose can affect all three foundational aspects of data analysis: sample size, demographic composition, and substantive outcomes. Thus, researchers' selection of ethnic classification method influences conclusions about inequities between ethnic groups, as well as subsequent decisions on how to address these. As stated in the conclusion of each of the individual studies, this indicates the importance for researchers to critically select their ethnic classification method based on their specific research question and context, and to be transparent about these decisions in publications.

## 7.2 Discussion of Ethnic Classification Methods

In order to assist researchers in making a decision on the most suitable ethnic classification method for their research, considerations for each classification method are discussed below, with a focus on integrating the findings from the three studies as well as the existing literature. I will also include insights and issues to be aware of from my experience in working with these ethnic classification methods throughout my PhD. It is important to emphasise that, if relevant in the research context (e.g., CANZUS states), Indigenous data sovereignty should take precedence—that is, when data include Indigenous Peoples, Indigenous Peoples have the right to determine what ethnic classification method to use (alongside all other decisions made throughout the research process; Cormack et al., 2019; Kukutai & Taylor, 2016). This is crucial given that (1) research often uses ethnicity data to the detriment of Indigenous Peoples, either intentionally or unintentionally (Smith, 2021; Walter

& Andersen, 2013); and (2) the three studies in this thesis show that ethnic classification method can influence research conclusions.

### 7.2.1 Total Response Grouping

Total response grouping, where respondents are counted in each of the ethnic groups they report, is one of two ethnic classification methods endorsed by Statistics New Zealand (2005). In samples with higher multi-ethnic prevalence, this results in total ethnic group counts that substantially exceeds the total number of participants. For example, Study 1 showed that the sum of total response ethnic groups was larger than the sample size by around 40% in children (38% multi-ethnic) and adolescents (32% multi-ethnic), compared to 14% in adults (16% multi-ethnic). In general, total response grouping is well-suited for research that needs to capture all the participants who identify with an ethnic grouping (e.g., the mental health and wellbeing of Pacific Peoples; Ataera-Minster & Trowland,s 2018). The method is relatively straightforward to implement in descriptive statistics, although it should be noted that, when used at a population level, it can increase the influence of the dominant ethnic group and decrease the influence of Indigenous and ethnic minority groups, making it unsuitable for purposes such as resource allocation. Total response grouping also counts multi-ethnic individuals multiple times, meaning that multi-ethnic participants are overrepresented in the data. In situations where multi-ethnic identification is associated with an outcome (e.g., adolescent mental health in Study 3), reporting outcomes by total response grouping risks inaccurate inferences about the prevalence of the outcome at the overall population level. Some of these concerns may be mitigated by (1) using Statistics New Zealand's (2005) standard

footnote for total response ethnicity output,[30] and (2) if relevant, clearly reporting an overall prevalence rate (using the total number of respondents as the denominator).

Total response grouping is also relatively simple to implement in inferential statistics when there is a single ethnic group of interest (either in isolation or in comparison to a reference group). However, when there are more than two ethnic groups of interest, total response grouping is more complex to implement because the outputted groups are not mutually exclusive—a characteristic typically needed for many common inferential techniques (e.g., chi-square test of independence, ANOVA, regression, and SEM; Cohen et al., 2003; Field et al., 2012). For example, as illustrated in Study 3, regression analyses using total response grouping typically involve a series of models containing one binary indicator at a time, resulting in five separate models instead of one. The results in Study 3 show that researchers' choice of reference group for the binary indicator can influence conclusions on ethnic group differences. In the context of adolescent health outcomes, modified total response typically resulted in larger between-group differences than original total response, although there were variations by ethnic group and the specific outcome measure. This emphasises the importance for researchers to critically select the reference group based on the purpose of the study. In general, if the effects of ethnic disadvantage and discrimination are of interest, modified total response (where the reference group only includes the privileged group) is recommended over original total response (where the reference group includes other marginalised ethnic groups; Lachowsky et al., 2020).

[30] Statistics New Zealand's (2005) standard footnote for total response ethnicity output is: "People who reported more than one ethnic group are counted once in each group reported. This means that the total number of responses for all ethnic groups can be greater than the total number of people who stated their ethnicities" (p. 6).

### 7.2.2 Sole/Combination Grouping

Sole/combination grouping, where respondents are counted once in the ethnic group or combination of ethnic groups they report, is the other ethnic classification method endorsed by Statistics New Zealand (2005). This method allows for more nuanced analyses when specific ethnic combinations are used (e.g., Māori/European, Pacific/European, Māori/Pacific/European, etc.; rather than an aggregate "multi-ethnic" category). For example, in Study 2, analysis by ethnic combination showed that Asian/European children (by mother proxy) and adolescents were more likely to not select a main ethnic group when compared to other ethnic combinations. In addition, analysis by sole/combination grouping in Study 3 showed that combination groupings tended to have poorer mental health outcomes than their constituent ethnic groups. However, this level of nuance means that results delineated by sole/combination grouping need to be interpreted and communicated very carefully, so that specific groups are not stigmatised. For example, it was important in Study 3 to situate the tendency of multi-ethnic participants to have poorer mental health in the context of negative societal stigma and discrimination towards multi-ethnic individuals (Sanchez et al., 2020; Skinner et al., 2020). Moreover, although sole/combination grouping is frequently used in U.S. research (Charmaraman et al., 2014), it is less common in Aotearoa New Zealand for two main reasons: (1) it can make Māori less visible in results, as some respondents who identify as Māori may be allocated to aggregated combination groupings; and (2) combination groupings evoke parallels with notions of "blood quantum" (e.g., "half-caste" and "quarter-caste"), which was historically used by the state to measure assimilation of Māori (Cormack & Robson, 2010). These ethical concerns need to be considered particularly in research situated in Aotearoa New Zealand.

From a statistical perspective, on one hand, sole/combination grouping can be attractive because the outputted categories are mutually exclusive. On the other hand, specificity in ethnic

combinations results in a large number of categories, making it unwieldy especially when exploring interaction effects between ethnicity and other variables. As illustrated in Study 2, adding the interaction between ethnic combination and socioeconomic deprivation resulted in an additional 14 terms in the adolescent regression model, thus reducing the events per variable and increasing the risk of biased results (Vittinghoff & McCulloch, 2007). Specific ethnic combinations also result in small subgroup sizes for some combinations, and unequal explanatory power between groups. For example, in the Youth'12 sample with 8,500 participants, 3,989 (47%) identified solely with a European ethnic group, whereas only 78 (1%) identified as both Māori and Pacific (see Study 1). Particularly in smaller datasets, less common ethnic combinations would have to be aggregated into an "other ethnic combinations" category for statistical analysis, thus limiting the level of insight provided. Finally, the specificity of sole/combination grouping means that group membership is more likely to fluctuate over time than, for example, total response grouping or administrative-prioritisation (Didham, 2005). This issue needs to be considered particularly in time series analyses.

### 7.2.3 Administrative-Prioritisation

Administrative-prioritisation, where respondents who report multiple ethnic groups are assigned to a single category according to a predetermined hierarchy, was discontinued as a standard ethnic classification method in Aotearoa New Zealand in 2005 (Statistics New Zealand, 2005). However, this method remains widely used in the health and education sectors, particularly when administrative datasets (rather than primary datasets) are involved (Education Counts, 2014; Ministry of Health, 2017). In addition, from my consultation with Māori researchers, administrative-prioritisation appears to be the preferred method for comparing more than two ethnic groups, because it prioritises Māori in recognition of *Te Tiriti o Waitangi*. This highlights the incongruity that can occur between official statistical standards and tangata whenua (people of the land). On the other hand, there are concerns that

administrative-prioritisation undercounts Pacific Peoples—another group that is of policy interest in Aotearoa New Zealand, and one that also tends to have relatively high rates of multi-ethnic identification (Statistics New Zealand, 2020a). Study 1 found that, compared to total response, administrative-prioritisation decreased the outputted Pacific proportion by up to 5% (5% in children, 3% in adolescents, 1% in adults). While this may appear small, the effect will likely increase as multi-ethnic identification increases.

In scenarios where there is a specific group of interest, one possible option for researchers is to adapt the standard prioritisation hierarchy, if considered appropriate. For example, if a combination grouping (e.g., Māori/Pacific) is of interest, one possibility is to use a "hybrid" approach that places the combination grouping at the top of the hierarchy (e.g., Māori/Pacific > all other Māori > all other Pacific > Asian > MELAA > Other > European). Irrespective of the specific hierarchy used, researchers need to be aware that with administrative-prioritisation, groups higher in the hierarchy will include a larger proportion of multi-ethnic participants than groups lower in the hierarchy. As discussed in Study 3, this means multi-ethnic identification can be a "confounding" factor in reported ethnic disparities. In other words, it is possible that observed ethnic differences are partially attributable to contextual factors that are detrimental to multi-ethnic individuals, such as multi-ethnic stigma and discrimination (Sanchez et al., 2020; Skinner et al., 2020).

Although use of the standard administrative-prioritisation algorithm can stem from the aim of elevating Māori in recognition of *Te Tiriti o Waitangi*, from a statistical perspective, administrative-prioritisation is also ostensibly popular in population studies because it outputs multiple ethnic identifications into a small number of mutually exclusive categories, without putting the onus on multi-ethnic individuals (e.g., by asking for their "main" ethnicity in a follow-up question). However, it is important to note that, consistent with previous research (Atatoa Carr et al., 2017; Kukutai & Callister, 2009), the current thesis found administrative-

prioritisation and self-prioritisation to be poorly correlated. Specifically, in each age cohort examined, there were discrepancies between the two prioritisation methods over 60% of the time (Study 1), in a way that was systematically associated with contextual characteristics (e.g., neighbourhood ethnic composition and socioeconomic deprivation; Study 2). Therefore, while administrative-prioritisation may be appropriate especially when certain ethnic groupings need to be elevated for equity reasons, researchers need to be aware that administratively-prioritised ethnicity often is not the ethnicity that multi-ethnic participants identify with most.

### 7.2.4 Self-Prioritisation

Due to the discrepancies between administrative-prioritisation and self-prioritisation, self-prioritisation may initially appear to be an ideal option if researchers want to classify multi-ethnic identifications into a small number of mutually exclusive categories, because it asks multi-ethnic participants to self-select their main ethnicity. However, the self-prioritisation question as typically worded (e.g., "Which is your main ethnic group?") makes two assumptions. First, the question assumes that multi-ethnic participants have a main ethnicity, which may not necessarily be the case. Study 1 shows that, when given a number of alternative response options (e.g., select two main ethnic groups, don't know, or refuse response), around 20% of multi-ethnic participants opted not to select a single main ethnic group. Therefore, as argued in Section 5.5.4, in scenarios where it is important to elicit the main ethnic group of participants who have one, it is recommended that the self-prioritisation question is modified to "Please indicate your main ethnic group(s) if you have one". This avoids portraying that having a main ethnic group is "normative" and "desirable", and that multi-ethnic identification is "unvalued". "Non-responses" to this question should be treated as a legitimate response (e.g., a declaration of multi-ethnic identity or a rejection of traditional ethnic categories; Aspinall & Song, 2013), and should not be omitted from the analytic sample.

Second, the self-prioritisation question assumes that multi-ethnic participants are free to choose the ethnicity they identify with most. However, consistent with previous research (e.g., Borrell, 2005; Herman, 2004), results from Study 2 suggest that responses are likely to be influenced by societal stereotypes, meaning that participants' responses may not necessarily reflect their private self-identification. Therefore, researchers should reflect on what they are interested in measuring via the self-prioritisation question. In many cases, researchers may actually be using self-prioritisation as a proxy for ethnic identity (Roth, 2016). If this is the case, a direct measure of ethnic identity (e.g., MEIM; Phinney & Ong, 2007) is recommended. However, if the purpose is to capture the ethnic group that multi-ethnic individuals most strongly affiliate with if they have one, self-prioritisation may be useful, provided that researchers consider the ethical concerns discussed above.

## 7.3 Limitations and Future Research

While this thesis advances knowledge and understanding of the effects of ethnic classification method on quantitative research in increasingly multi-ethnic contexts, some limitations related to the overall research need to be noted (specific limitations related to the individual studies are discussed in their respective chapters). Directions for future research are also outlined. First, while the use of large-scale child, adolescent, and adult datasets is a strength of this thesis, there are some limitations with these data. Specifically, although the child and adolescent datasets are broadly nationally representative of their age cohort at the time (Clark et al., 2013; Morton et al., 2015), the adult datasets used in Study 1 (mothers and partners) and Study 2 (mothers only) are not representative of the wider adult population, and thus may not be generalisable to older adults in particular. In addition, the data are somewhat dated (adult data were collected in 2010, adolescent data in 2012, and child data in 2015), especially given that (multi-)ethnic identification is a dynamic social phenomenon highly

influenced by changes in societal conceptions and perceptions of ethnicity (Aspinall, 2018b). It would be interesting to examine the similarities and/or differences in ethnic identification patterns between the 2010s and 2020s. Also, ethnicity data for children in this thesis were parent-identified rather than self-identified. While this is a common way of collecting ethnicity data for children (Perez, 2006; Statistics New Zealand, 2005), it will be interesting to compare this to child-identified ethnicity, which was collected by Growing Up in New Zealand for the first time in the eight-year data collection wave.

Second, there are two key limitations with the way ethnicity was operationalised in this thesis. First, ethnic identification was measured cross-sectionally, which can give the false impression that ethnicity is a fixed characteristic. In reality, ethnic identification is fluid across time and contexts (Carter et al., 2009; Didham, 2016; D. R. Harris & Sim, 2002; Liebler et al., 2017; Nishina et al., 2010), so the ethnicity data utilised in this thesis can only be interpreted as participants' responses in the time and context the survey was administered. However, cross-sectional examination of ethnic identification is still valuable, especially given that this is how ethnicity data are most commonly collected and analysed. Second, this thesis examined the effects of ethnic classification method on broad ethnic groupings. While this is a common way of outputting ethnicity data for quantitative analysis, there is considerable heterogeneity within each ethnic grouping. Future research can explore how researchers' classification of ethnicity data into broad versus specific ethnic groups impact on the conclusions drawn from analyses (e.g., Peiris-John et al., 2022). However, it should be noted that, just as the use of broad ethnic groupings can be problematic, using specific ethnic groups can also have limitations. For example, depending on the sample, some specific ethnic groups may have subgroup sizes that are too small for meaningful statistical analysis (Gillborn et al., 2018).

Third, there are limitations related to the variables used alongside ethnicity in this thesis. Many of these variables are socially constructed and heavily influenced by a Western

worldview (e.g., socioeconomic deprivation, urbanicity, and mental health). Like ethnicity, there are many ways to measure these variables, and these decisions can have an impact on results. The focus of this thesis was to examine the effects of ethnic classification method while keeping the measurement of other variables constant. However, it would be worthwhile for future research to explore how researchers' decisions on the measurement of other variables, as well as the interaction between how ethnicity and other variables are measured, impact on data analysis and research conclusions (e.g., Hobbs et al., 2019; Salmond & Crampton, 2012).

Finally, this thesis is solely based on secondary quantitative data, and the implications for methodological decisions related to ethnic classification are based on my interpretations (with input from my supervisors, advisors, co-authors, and colleagues). It would be valuable for future research to directly interview stakeholders, particularly those belonging to Indigenous and/or ethnic minority groups, on their perspectives on how ethnicity data should be categorised in quantitative research. In addition, it would be valuable to directly ask participants for their thoughts on how ethnicity data are collected and outputted. One possibility is conducting cognitive interviews to explore how participants interpret and respond to questions about the ethnic group(s) they identify with (i.e., total response), as well as their "main" ethnic group (i.e., self-prioritisation).

## 7.4 Implications

Irrespective of these limitations, this thesis has important implications for ethnic classification in the increasingly multi-ethnic global context. Most crucially, the results empirically show that research involving ethnicity data is inherently political and subjective, and that researchers are in a position of power, because their decisions on ethnic classification method can influence quantitative analyses—from outputted ethnic group sizes (Study 1), to the demographic composition of these outputted ethnic groups (Study 2), to the substantive

conclusions drawn (Study 3). Under a critical research paradigm, it is not enough to simply document the effects that different ethnic classification methods have on quantitative research. Rather, these findings need to result in positive social change, particularly in terms of equity, social justice, and emancipation for those who are oppressed by power (Gillborn et al., 2018; Lincoln et al., 2018; Stage, 2007; Stage & Wells, 2014). Through this lens, the implications of this thesis are discussed below for two groups of people: producers of research (e.g., researchers), and audiences of research (e.g., policymakers, practitioners, university students, the media, and the general public).

### 7.4.1 For Researchers

For quantitative researchers working with ethnicity data, it is first and foremost paramount that they critically select the most appropriate ethnic classification method for their research, and are transparent about this decision and its implications. This is because this thesis empirically shows that researchers' decisions can influence the knowledge that is produced, and hence the implications for policy and practice. It is difficult to give specific recommendations about which ethnic classification method is "best", as each method has strengths and limitations. Rather, as stated in the discussion section of the three individual studies, the most appropriate method depends on the research question(s) and context. To summarise, total response grouping is well-suited for scenarios where all the members of an ethnic group need to be captured, sole/combination grouping is helpful for adding nuance to analyses, administrative-prioritisation is advantageous when certain ethnic group(s) need to be elevated for equity purposes, and self-prioritisation—if asked sensitively—may be useful if strength of self-affiliation is needed. From a pragmatic perspective, decisions also needs to be considered in light of the properties of the analytic dataset (e.g., sample size). More detailed considerations for each ethnic classification method have been discussed above (see Section 7.2).

Similar to selecting variables for statistical models, it is important that decisions regarding ethnic classification method are made a priori based on theory (Heinze et al., 2018). In other words, the ethnic classification method should not be selected a posteriori based on results that best fits the research agenda. However, sensitivity analyses can be useful for understanding the effects that different ethnic classification methods may have on research findings. As discussed previously, in research contexts where this is relevant (e.g., CANZUS states), it is crucial that decisions around ethnic classification method (and indeed all decisions made during the research process, see Table 7.1) involve consultation with Indigenous researchers and/or stakeholders, as well as researchers and/or stakeholders of other ethnic minority groups, because these decisions can have both direct and indirect repercussions on marginalised communities (Cormack et al., 2019; Hudson et al., 2010; Kukutai & Taylor, 2016; Research Data Alliance International Indigenous Data Sovereignty Interest Group, 2019).

In addition to implications specific to ethnic classification decisions, the subjective nature of ethnicity as clearly evidenced in this thesis also have important implications on the broader research process. Some key questions to consider at each stage of the research process when working quantitatively with ethnicity data is outlined in Table 7.1, and briefly discussed below.

1. *Research conceptualisation*: from the outset of developing the research topic and question(s), it is important for researchers to consider whether ethnicity is an important variable to include. For example, in Study 3's investigation of adolescent mental health outcomes, ethnic-specific analyses provided insight into ethnic groups that need additional support, which would have been missed if the analyses only focused on other characteristics (e.g., sex and socioeconomic deprivation). If ethnicity is a variable of interest, consultation with cultural advisors should occur—both here and throughout the research process—to ensure that the research aligns with the interests of, and will bring benefit to, the Indigenous

**Table 7.1**

*Key Questions in the Research Process When Working Quantitatively with Ethnicity Data*

| Research Stage | Key Question(s) |
|---|---|
| 1. Research conceptualisation | Is ethnicity an important variable to include? Why or why not? |
| 2. Research design | How does my positionality influence the research design? |
| 3. Operationalisation of concepts | How will ethnicity be defined and measured? |
| 4. Data collection | How will ethnicity data be collected? Is the sample size adequate? |
| 5. Data processing | How will I classify (i.e., output) ethnicity data for analysis? |
| 6. Data analysis | How does my chosen ethnic classification method potentially influence results? What other variables do I need to include alongside ethnicity? |
| 7. Interpretation | Have I situated the results within the wider sociocultural context? |
| 8. Communication | Have I been transparent about my decisions regarding ethnicity measurement? |

*Note.* Research process adapted from Punch (2014).

and ethnic minority groups involved.

2. *Research design*: given the subjectivity of research involving ethnicity, it is important for researchers to be reflexive on how their positionality influences their research design, including what data are collected (e.g., what communities are represented, what variables are prioritised, etc.), and how these data are collected (e.g., in person, online, etc.).
3. *Operationalisation of concepts*: careful consideration is needed around how ethnicity is defined and measured in the research at hand, as this influences what ethnicity data are collected, and hence, what data are available for analysis. For example, if Statistics New Zealand's (2005) definition of ethnicity is used (see Section 2.4.2), ethnicity is likely to be measured through participants' self-identification with each their ethnic group(s).

Depending on the research purpose and context, it may be appropriate to measure ethnicity in alternative ways, such as through self-prioritised ethnicity (when strength of self-affiliation is of interest) or socially-ascribed ethnicity (when external perceptions of one's ethnicity is of interest, e.g., for studies on discrimination; Roth, 2016; White et al., 2020).

4. *Data collection*: the ethnicity data that are collected need to match the way ethnicity is operationalised in the study. The response options that are provided also need to be considered—including, if relevant, what specific ethnic groups are listed on the form, how the open-ended response option is formatted, and what alternative options are given for the self-prioritisation question (Studies 1 and 2 in this thesis suggest that more restrictive alternative options may result in a higher proportion of multi-ethnic participants who choose a "main" ethnic group, but as discussed, this needs to be balanced with ethical concerns). Researchers also need to ensure that they have an adequate sample size—both overall and by the specific ethnic group(s) of interest—to allow meaningful ethnic-specific analyses.
5. *Data processing*: the collected ethnicity data need to be classified (i.e., outputted) into a format suitable for data analysis. Two main decisions include: (1) whether specific ethnic groups are aggregated into broader ethnic groupings, and if so, what groupings are used (e.g., European, Māori, Pacific, Asian, MELAA, and Other); and (2) if total response ethnicity was collected, how multiple responses will be outputted (e.g., total response grouping, sole/combination grouping, or administrative-prioritisation). The data will need to be processed according to these decisions (e.g., if administrative-prioritisation is selected, researchers will need to run an external prioritisation algorithm to output total response ethnicity into administratively-prioritised ethnicity). This thesis empirically shows that these decisions can affect research findings and conclusions.

6. *Data analysis*: during data analysis, it is important to consider how the chosen ethnic classification method may affect results. Researchers may like to consider conducting sensitivity analyses to empirically explore this. In addition, researchers need to consider what other variables to include in the analysis (e.g., sex/gender, socioeconomic deprivation, ethnic identity, perceived ethnic discrimination, etc.), and whether interactions between ethnicity and these variables need to be explored to gain insight into how results differ for each ethnic group.
7. *Interpretation*: while all steps in the research process are influenced by the researcher, interpretation of results is arguably where researcher influence is the most substantial (Bonilla-Silva & Zuberi, 2008). This is because the same data and results can be interpreted differently by different people depending on their positionality and worldview. When interpreting ethnicity data, it is critical that this is situated within the wider sociocultural context (e.g., interpersonal and structural racism), so that results associated with ethnicity are not attributed to the individuals themselves (e.g., biological essentialism; Gillborn et al., 2018; Zuberi, 2001).
8. *Communication*: finally, it is important that decisions around ethnicity measurement (e.g., the ethnic classification method selected) are made transparent in research dissemination (e.g., publications and presentations). Transparency is crucial both for scientific replicability, as well as for raising awareness of the complexities around ethnic classification. There is some evidence that journal editorial guidelines on the reporting of ethnicity can help improve reporting practices (Sankar et al., 2015).

The list above focuses on specific considerations at each step of the research process. These considerations can ultimately be summed up in a short but imperative overarching question: *who will benefit from the research?* Critical reflection of this question throughout the research process is crucial given the potential for research involving ethnicity data to

emancipate, or further oppress, marginalised ethnic groups (Smith, 2021; Walter & Andersen, 2013).

### 7.4.2 For Research Audiences

This thesis also has important implications for research audiences, particularly around raising awareness of the complexities of ethnic classification, as well as the subjective nature of quantitative research. For policymakers and practitioners (e.g., healthcare professionals, social workers, etc.), it will be valuable to inform them of the impact that researchers' ethnic classification choice can have on quantitative results via outputted ethnic group sizes (Study 1), the demographic composition of these ethnic groups (Study 2), and substantive findings regarding ethnic disparities in outcomes (Study 3). While the onus should be on researchers to select the most appropriate ethnic classification method and make this transparent, policymakers' and practitioners' understanding of these methodological issues will help them read and translate research into practice in a more critical way, and question research that does not adequately address these issues.

This thesis also shows that it will be important to educate university students—especially those doing health or social science courses with a research component—on the history of race and ethnicity, as well as the subjectivity of ethnicity measurement and quantitative research more generally. Contrary to (post-)positivism—the dominant paradigm underpinning quantitative research (Lincoln et al., 2018)—the current thesis demonstrates that research involving ethnicity data cannot be conducted in a neutral manner. Rather, research involving ethnicity data is inherently subjective and political, and, as this thesis shows, researchers' methodological decisions can influence the results and conclusions drawn. It would be valuable for the next generation of potential researchers to be cognisant of these issues before engaging in research, so that future research is increasingly produced from a

critical perspective that aims to reverse the harms on marginalised ethnic groups, towards equity and social justice for all.

Finally, journalists should be informed of the effects that decisions on ethnicity measurement can have on news stories, particularly because these often include reporting of ethnic disparities and can contribute to negative societal stereotypes. Using the results from Study 3 as an example, a news story could have either reported that Māori youth have significantly higher rates of self-harm than European youth, or that there are no significant differences between the two ethnic groups, depending on which ethnic classification method was used to analyse the data (i.e., administrative-prioritisation and self-prioritisation, respectively). It is therefore imperative that news stories involving ethnic disparities are reported in a transparent, nuanced, and socially responsible manner. It may be helpful for longform journalism to include stories on the impact that researchers' methodological decisions can have on substantive research outcomes, to help the general public engage in media and science more critically. One possibility is an accessible longform article on the key results and implications of Study 3, which investigated the effects that researchers' ethnic classification decisions can have on substantive outcomes in adolescent mental health.

## 7.5 Contributions

This thesis was precipitated by the questions of how to choose the most appropriate ethnic classification method to use in my research in an adolescent sample with over 30% multi-ethnic identification, and how this choice would affect my results. In my search for answers in the existing literature, I discovered some relevant theoretical pieces (e.g., Cormack & Robson, 2010; Denton & Deane, 2010; OMB, 2000), as well as some empirical studies which explored the effects of ethnic classification method in datasets with relatively low multi-ethnic prevalence (<10%; Boven et al., 2020; Liebler & Halpern-Manners, 2008; Ministry of Health,

2008; Rutkowski et al., 2017; Te Rōpū Rangahau Hauora a Eru Pōmare, 2000). However, there were two main gaps in the literature: (1) there was a lack of empirical studies on ethnic classification effects in datasets with higher levels of multi-ethnic prevalence; and (2) there appeared to be a disconnect between the literature and mainstream research practice, because although the literature emphasised the importance of ethnic classification choice, many researchers were not adequately transparent about their decisions in this area (Moubarac, 2013; Nishina & Witkow, 2020). The current thesis aimed to address these gaps using a critical quantitative approach.

In terms of addressing the first gap, this thesis—through three peer-reviewed journal articles—contributes comprehensive empirical information on the effects of common ethnic classification methods in increasingly multi-ethnic age cohorts (adults = 16% multi-ethnic; adolescents = 32% multi-ethnic; children = 38% multi-ethnic). These effects were explored in three important areas of quantitative analysis: ethnic group size (Study 1), the demographic composition of ethnic groups (Study 2), and substantive findings on ethnic disparities in outcomes (Study 3). This comprehensive investigation using large-scale datasets means that the current results will likely be relevant to researchers irrespective of the rate of multi-ethnic identification in their sample, the ethnic classification method they choose to use, and the specific academic discipline they are in. Through this investigation, the three studies provide strong empirical support to the existing literature on the importance of researcher criticality and transparency regarding ethnic classification. In addition, this thesis makes two other particularly noteworthy contributions to the literature. First, response patterns to self-prioritised ethnicity, as well as its effects on quantitative analysis, is comprehensively examined in three age cohorts (previous research has typically focused less on self-prioritisation, despite its popularity in applied research). Second, investigation into the effects of ethnic classification method was extended to the field of mental health—a field that is arguably more sensitive to

issues of ethnic identification and identity (previous research has typically focused on ethnic classification effects on physical health outcomes).

In terms of addressing the disconnect between ethnic classification literature and mainstream research practice, the current thesis provides unequivocal empirical evidence that researchers' decisions regarding ethnic classification impact quantitative results. This highlights the importance for researchers to critically engage in issues regarding the subjectivity of ethnicity measurement. Many researchers who are not fully aware of these issues may, like me at the beginning of my PhD, contend that ethnic classification methods only warrant attention if there is sufficient empirical evidence that they impact research outcomes. Based on my experience, evidence that ethnic classification method empirically matters contributes to bridging the disconnect between the literature and mainstream practice, because it opens up dialogue as to why criticality in ethnicity measurement is imperative, not just from a statistical perspective, but also from theoretical, ethical, equity, and Indigenous data sovereignty perspectives. Opportunities to share my PhD research, both formally (e.g., in seminars) and informally (e.g., through conversations with peers and colleagues), have resulted in people expressing that it prompted them to think more deeply and critically about ethnicity measurement, and in particular, the limitations of self-prioritised ethnicity.

Moreover, this thesis has contributed a unique perspective to the field of ethnic classification. Specifically, it documents my journey as an Asian person—often considered an outsider in ethnic classification issues in Aotearoa New Zealand—as I embarked, without any preconceived agenda or allegiance, on an "objective" search for which ethnic classification method to use in my research (I initially thought objectivity and neutrality was possible). It was only on this journey that I truly realised the importance of being—and took on the positions of—a tangata Tiriti, an ally of Indigenous sovereignty, and a critical quantitative researcher. It is my hope that this thesis will also inspire others onto similar journeys.

CHAPTER 8

# Conclusion

Ethnicity is an important variable in research and policy for equity and social justice, but measurement of ethnicity is complex due to changing patterns in self-identified ethnicity, including steadily increasing multi-ethnic identification rates (Aspinall, 2018b; Atatoa Carr et al., 2017; Balestra & Fleischer, 2018; N. Jones et al., 2021). Existing literature suggests that researchers' ethnic classification decisions can influence research outcomes (Boven et al., 2020; Callister et al., 2007; Mays et al., 2003; Rutkowski et al., 2017). However, there is a lack of empirical research on the effects that different ethnic classification methods have on quantitative research in samples with higher multi-ethnic prevalence (e.g., >10%). In addition, many mainstream researchers appear to be unaware of the complexities involved in the operationalisation of ethnicity (Moubarac, 2013).

In this thesis, I address these gaps using a critical quantitative approach from my positionality as an Asian tangata Tiriti. Specifically, I utilised large-scale datasets from increasingly multi-ethnic age cohorts (adults, adolescents, and children) to examine the effects that common ethnic classification methods have on three important areas in quantitative analysis: outputted ethnic group size, the demographic composition of these ethnic groups, and substantive outcomes. The corresponding three studies showed that researchers' decisions on ethnic classification can have a considerable impact on each of these areas, highlighting the importance for researchers to be critical and transparent when selecting an ethnic classification method. The subjective and political nature of research involving ethnicity data also has important implications for the research process more broadly, particularly around researcher criticality and reflexivity.

To summarise, the central argument of this thesis is as follows:

- *Ethnicity matters*—because it is associated with widespread inequities as a result of historical and ongoing racism;
- *Positionality matters*—because researchers' positionality influences each step of the research process;
- *Method matters*—because ethnic classification method affects outputted ethnic group size, demographic composition, and substantive outcomes;
- *Context matters*—because it can affect ethnic self-identification, especially when self-prioritisation is involved;
- *Outcome matters*—because the impact of ethnic classification method can differ depending on the specific outcome measure; and
- *Interpretation matters*—because research involving ethnicity is not objective or neutral, but inherently subjectively and political, and has both direct and indirect implications on equity and social justice for oppressed groups.

More succinctly, researcher criticality and reflexivity are imperative in research involving ethnicity, because researchers are in a position of immense power. As this thesis unequivocally shows, researchers' methodological decisions influence data analysis, results, and conclusions; and ultimately, knowledge construction, policy, and practice.

# Afterword

I began this thesis with two methodological questions: (1) how I will choose what ethnic classification method to use in my research, and (2) how my chosen method will influence my results. I thought I would be able to quickly arrive at an objective answer, and then carry on with researching what I was truly interested in at the time: ethnocultural differences in adolescent depression and its implications for practice. The process turned out to be much more complex, and much more than just a methodological exercise. It took me on an academic and personal journey I never would have envisaged.

Academically, one of the most confronting learnings for me was that it is impossible for researchers to be completely neutral and objective. Instead, each step of the research process is subjective, political, and influenced by our positionality. I realised that engaging in research comes with great power, and with great power, immense responsibility—from the overarching decisions that I make, down to every word that I choose to use. Reflecting on my positionality made me frequently question my place as an Asian person—without lived multi-ethnic experience at that—researching (multi-)ethnic classification in bicultural Aotearoa New Zealand. No doubt internalised racism played a role in this recurrent intrusive thought: "Who are you to be exploring this topic? You random Asian should just go home to where you belong!" (except Aotearoa feels much more like home to me than Hong Kong). I was also aware that my "otherness" is immediately apparent to others, from my physical appearance through to my surname. I worried I would be perceived as illegitimate, and that the conclusions I drew from my results would be wrong due to my positionality. I am so thankful and privileged to have had valuable input and feedback from my supervisors, advisors, co-authors, colleagues, and peers throughout my PhD; as well as their reassurance that I can bring a unique perspective to this space as an Asian tangata Tiriti.

Beyond my own research, my realisation of the subjectivity of science made me question many things I once thought of as "truth". Initially, it was deeply disorienting to think that Western science is not the "holy grail of objective knowledge" I had grown up believing and valuing. Interestingly, the scientific response to the COVID-19 pandemic over the past two years helped me regain some faith in science (I write this during the March 2022 Omicron outbreak in Aotearoa New Zealand). I came to realise that the subjectivity of science does not render it useless; instead, it emphasises the importance of criticality, transparency, and engaging in science for the right reasons. I now view science with a much healthier dose of scepticism, and prioritise approaching research critically with the goals of equity and social justice. In this sense, I have gained renewed direction.

More personally, as previously alluded to, researching ethnic classification confronted me with who I am as a person. Prior to my PhD, I had—ironically—largely put my own ethnicity into a locked box at the back of my head, because it was too difficult to think that I was different and did not belong. In the past, whenever I was asked to indicate my ethnicity, I would quickly tick what I thought was expected of me (usually "Chinese") and move on, because it made me feel uncomfortable and "outed" as a perpetual foreigner despite all my attempts to fit in. During my PhD, I came to understand that another reason the ethnicity question made me feel uncomfortable was that, by identifying as Chinese, I was categorising myself into a culture that—while it has some beautiful aspects—has also brought me some personal hurt. Given the choice, I tended to prefer identifying as "Asian" because I felt that it sufficiently satisfies the "upper powers", while being broad enough to lack any real meaning. But more than that, I wished the ethnicity question was never asked.

Over the course of my PhD, I realised ethnicity data are crucial for addressing societal inequities. However, this also opened my eyes to the extent of racism entrenched in society (I previously dismissed experiences of racism as me being too sensitive). I became deeply

disillusioned with humanity—why can't we live in a world with no inequities and no racism, where ethnicity is no longer a relevant variable? Is that a place we can ever get to? Serendipitously, this disillusionment led me to engage with critical quantitative literature, and contributed to me shifting from a post-positivist paradigm to a critical quantitative paradigm. When I first changed my PhD topic from adolescent depression to ethnic classification, I admittedly felt sadness at having to "let go" of my interest in depression research. In hindsight, I have not had to let it go. In many ways, addressing issues of racism and social injustice is, in fact, addressing mental health challenges at its root.

Coming back to the initial question of which ethnic classification method is best to use: on one hand, part of me still wishes I could have objectively arrived at a universal best method. On the other hand, I think there is beauty in the fact that there is no single best method, because that reflects the nature of ethnicity, as well as humanity more broadly. Although there is value in ethnic categorisation for equity and social justice, by nature we cannot be neatly fit into boxes; and through this journey, I am learning that I am comprised of more than just ticks (or lack thereof) in socially constructed categories. People are more than ticks in boxes—behind each tick is a person with rich lived experiences. And behind each thesis—including quantitative ones—is a person with their rich lived experience, how it has shaped their research, and equally, how their research has shaped them. This has been a glimpse into the deeply personal journey this PhD has taken me on.

APPENDIX A

# Supplementary Tables for Study 1

**Supplementary Table A.1**

*Two-Sample z-Tests of Pairwise Age Differences in Dual- and Multi-Ethnic Participants' Selection Rate of One Main Ethnic Group*

| Group A | | | | Group B | | | | Group A vs. Group B | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Age Group | *N* | *n* | Row % | Age Group | *N* | *n* | Row % | % Difference (A-B) | *z* | *p* | *ES*[a] |
| Children[b] | 2359 | 1814 | 77 | Adolescents | 2712 | 2601 | 96 | -19 | -20.12 | <.001 | -0.60 |
| Children[b] | 2359 | 1814 | 77 | Adults | 1758 | 1410 | 80 | -3 | -2.55 | .011 | -0.08 |
| Adolescents | 2712 | 2601 | 96 | Adults | 1758 | 1410 | 80 | 16 | 16.89 | <.001 | 0.51 |

[a]Cohen's (1988) *h* effect size.

[b]Child responses were collected by proxy from the child's mother.

**Supplementary Table A.2**

*Dual- and Multi-Ethnic Participants' Self-Selection of a Main Ethnic Group by Age Group and Ethnic Combination*

| | | Selected *one* main ethnic group | | | | Did *not* select one main ethnic group | |
|---|---|---|---|---|---|---|---|
| | | Matched admin-prioritisation | | Did *not* match admin-prioritisation | | | |
| | *N* | *n* | Row % | *n* | Row % | *n* | Row % |
| *Children*[a] *(N = 2,359)* | | | | | | | |
| Māori/European | 725 | 167 | 23 | 366 | 50 | 192 | 26 |
| Pacific/European | 195 | 69 | 35 | 84 | 43 | 42 | 22 |
| Asian/European | 161 | 13 | 8 | 100 | 62 | 48 | 30 |
| Other/European | 374 | 199 | 53 | 137 | 37 | 38 | 10 |
| Māori/Pacific | 144 | 58 | 40 | 41 | 28 | 45 | 31 |
| Other 2 ethnic groups | 247 | 107 | 43 | 80 | 32 | 60 | 24 |
| Māori/Pacific/European | 178 | 58 | 33 | 70 | 39 | 50 | 28 |
| Other 3+ ethnic groups | 335 | 61 | 18 | 204 | 61 | 70 | 21 |
| *Adolescents (N = 2,712)* | | | | | | | |
| Māori/European | 987 | 350 | 35 | 618 | 63 | 19 | 2 |
| Pacific/European | 393 | 225 | 57 | S | 41 | <10 | 2 |
| Asian/European | 234 | 100 | 43 | 115 | 49 | 19 | 8 |
| Other/European | 332 | 49 | 15 | 259 | 78 | 24 | 7 |
| Māori/Pacific | 78 | 47 | 60 | S | 37 | <10 | 3 |
| Other 2 ethnic groups | 238 | 142 | 60 | 82 | 35 | 14 | 6 |
| Māori/Pacific/European | 126 | S | 48 | 62 | 49 | <10 | 3 |
| Other 3+ ethnic groups | 325 | 147 | 45 | 156 | 48 | 22 | 7 |
| *Adults (N = 1,758)* | | | | | | | |
| Māori/European | 974 | 361 | 37 | 412 | 42 | 201 | 21 |
| Pacific/European | 217 | 87 | 40 | 78 | 36 | 52 | 24 |
| Asian/European | 61 | S | 21 | 40 | 66 | <10 | 13 |
| Other/European | 124 | 50 | 40 | 62 | 50 | 12 | 10 |
| Māori/Pacific | 109 | 51 | 47 | 33 | 30 | 25 | 23 |
| Other 2 ethnic groups | 111 | 67 | 60 | 29 | 26 | 15 | 14 |
| Māori/Pacific/European | 86 | 30 | 35 | 37 | 43 | 19 | 22 |
| Other 3+ ethnic groups | 76 | 18 | 24 | 42 | 55 | 16 | 21 |

*Note.* Cell counts less than 10 are suppressed as "<10". Secondary suppression (S) was applied to the next smallest cell in the row so that the suppressed cell cannot be recalculated.

[a]Child responses were collected by proxy from the child's mother.

**Supplementary Table A.3**

*Two-Sample z-Tests of Age Differences in the Discrepancy Rate between Administratively-Prioritised and Self-Prioritised Ethnicity of Dual- and Multi-Ethnic Participants*

| Group A | | | | Group B | | | | Group A vs. Group B | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Age Group | *N* | *n* | Row % | Age Group | *N* | *n* | Row % | % Difference (A-B) | *z* | *p* | *ES*[a] |
| Children[b] | 2359 | 1627 | 69 | Adolescents | 2712 | 1594 | 59 | 10 | 7.52 | <.001 | 0.21 |
| Children[b] | 2359 | 1627 | 69 | Adults | 1758 | 1081 | 61 | 7 | 5.00 | <.001 | 0.16 |
| Adolescents | 2712 | 1594 | 59 | Adults | 1758 | 1081 | 61 | -3 | -1.81 | .071 | -0.06 |

[a]Cohen's (1988) *h* effect size.

[b]Child responses were collected by proxy from the child's mother.

**Supplementary Table A.4**

*One-Sample z-Tests of Differences in Ethnic Group Proportions by Ethnic Classification Method—Children[a] (N = 5,604)*

| Group A (Reference Group) | | | Group B (Comparison Group) | | | Group A vs. Group B | | | |
|---|---|---|---|---|---|---|---|---|---|
| Classification Method | *n* | Row % | Classification Method | *n* | Row % | % Difference (A-B) | *z* | *p* | *ES*[b] |
| *European* | | | | | | | | | |
| Total response | 3727 | 67 | Self-prioritisation | 3016 | 54 | -13 | -20.12 | <.001 | -0.26 |
| Total response | 3727 | 67 | Admin-prioritisation | 2221 | 40 | -27 | -42.62 | <.001 | -0.55 |
| Self-prioritisation | 3016 | 54 | Admin-prioritisation | 2221 | 40 | -14 | -21.30 | <.001 | -0.29 |
| *Māori* | | | | | | | | | |
| Total response | 1165 | 21 | Self-prioritisation | 519 | 9 | -12 | -21.27 | <.001 | -0.33 |
| Total response | 1165 | 21 | Admin-prioritisation | 1165 | 21 | - | - | - | - |
| Self-prioritisation | 519 | 9 | Admin-prioritisation | 1165 | 21 | 12 | 29.77 | <.001 | 0.33 |
| *Pacific* | | | | | | | | | |
| Total response | 1077 | 19 | Self-prioritisation | 734 | 13 | -6 | -11.63 | <.001 | -0.17 |
| Total response | 1077 | 19 | Admin-prioritisation | 795 | 14 | -5 | -9.56 | <.001 | -0.14 |
| Self-prioritisation | 734 | 13 | Admin-prioritisation | 795 | 14 | 1 | 2.42 | .016 | 0.03 |
| *Asian* | | | | | | | | | |
| Total response | 914 | 16 | Self-prioritisation | 674 | 12 | -4 | -8.68 | <.001 | -0.12 |
| Total response | 914 | 16 | Admin-prioritisation | 821 | 15 | -2 | -3.36 | <.001 | -0.05 |
| Self-prioritisation | 674 | 12 | Admin-prioritisation | 821 | 15 | 3 | 6.04 | <.001 | 0.08 |
| *Other* | | | | | | | | | |
| Total response | 979 | 17 | Self-prioritisation | 661 | 12 | -6 | -11.19 | <.001 | -0.16 |
| Total response | 979 | 17 | Admin-prioritisation | 602 | 11 | -7 | -13.26 | <.001 | -0.19 |
| Self-prioritisation | 661 | 12 | Admin-prioritisation | 602 | 11 | -1 | -2.44 | .015 | -0.03 |

[a]Child responses were collected by proxy from the child's mother.

[b]Cohen's (1988) *h* effect size.

**Supplementary Table A.5**

*One-Sample z-Tests of Differences in Ethnic Group Proportions by Ethnic Classification Method—Adolescents (N = 8,354)*

| Group A (Reference Group) | | | Group B (Comparison Group) | | | Group A vs. Group B | | | |
|---|---|---|---|---|---|---|---|---|---|
| Classification Method | *n* | Row % | Classification Method | *n* | Row % | % Difference (A-B) | *z* | *p* | *ES*[a] |
| *European* | | | | | | | | | |
| Total response | 6259 | 75 | Self-prioritisation | 5293 | 63 | -12 | -24.38 | <.001 | -0.25 |
| Total response | 6259 | 75 | Admin-prioritisation | 3989 | 48 | -27 | -57.30 | <.001 | -0.57 |
| Self-prioritisation | 5293 | 63 | Admin-prioritisation | 3989 | 48 | -16 | -29.61 | <.001 | -0.32 |
| *Māori* | | | | | | | | | |
| Total response | 1669 | 20 | Self-prioritisation | 855 | 10 | -10 | -22.27 | <.001 | -0.28 |
| Total response | 1669 | 20 | Admin-prioritisation | 1669 | 20 | - | - | - | - |
| Self-prioritisation | 855 | 10 | Admin-prioritisation | 1669 | 20 | 10 | 29.38 | <.001 | 0.28 |
| *Pacific* | | | | | | | | | |
| Total response | 1423 | 17 | Self-prioritisation | 981 | 12 | -5 | -12.86 | <.001 | -0.15 |
| Total response | 1423 | 17 | Admin-prioritisation | 1181 | 14 | -3 | -7.04 | <.001 | -0.08 |
| Self-prioritisation | 981 | 12 | Admin-prioritisation | 1181 | 14 | 2 | 6.80 | <.001 | 0.07 |
| *Asian* | | | | | | | | | |
| Total response | 1262 | 15 | Self-prioritisation | 980 | 12 | -3 | -8.62 | <.001 | -0.10 |
| Total response | 1262 | 15 | Admin-prioritisation | 1044 | 12 | -3 | -6.66 | <.001 | -0.08 |
| Self-prioritisation | 980 | 12 | Admin-prioritisation | 1044 | 12 | 1 | 2.18 | .030 | 0.02 |
| *Other* | | | | | | | | | |
| Total response | 812 | 10 | Self-prioritisation | 245 | 3 | -7 | -20.94 | <.001 | -0.29 |
| Total response | 812 | 10 | Admin-prioritisation | 470 | 6 | -4 | -12.63 | <.001 | -0.16 |
| Self-prioritisation | 245 | 3 | Admin-prioritisation | 470 | 6 | 3 | 14.59 | <.001 | 0.13 |

[a]Cohen's (1988) *h* effect size.

**Supplementary Table A.6**

*One-Sample z-Tests of Differences in Ethnic Group Proportions by Ethnic Classification Method—Adults (N = 10,862)*

| Group A (Reference Group) | | | Group B (Comparison Group) | | | Group A vs. Group B | | | |
|---|---|---|---|---|---|---|---|---|---|
| Classification Method | *n* | Row % | Classification Method | *n* | Row % | % Difference (A-B) | *z* | *p* | *ES*[a] |
| *European* | | | | | | | | | |
| Total response | 6865 | 63 | Self-prioritisation | 6266 | 58 | -6 | -11.92 | <.001 | -0.11 |
| Total response | 6865 | 63 | Admin-prioritisation | 5639 | 52 | -11 | -24.39 | <.001 | -0.23 |
| Self-prioritisation | 6266 | 58 | Admin-prioritisation | 5639 | 52 | -6 | -12.18 | <.001 | -0.12 |
| *Māori* | | | | | | | | | |
| Total response | 1649 | 15 | Self-prioritisation | 1134 | 10 | -5 | -13.77 | <.001 | -0.14 |
| Total response | 1649 | 15 | Admin-prioritisation | 1649 | 15 | - | - | - | - |
| Self-prioritisation | 1134 | 10 | Admin-prioritisation | 1649 | 15 | 5 | 16.16 | <.001 | 0.14 |
| *Pacific* | | | | | | | | | |
| Total response | 1639 | 15 | Self-prioritisation | 1436 | 13 | -2 | -5.44 | <.001 | -0.05 |
| Total response | 1639 | 15 | Admin-prioritisation | 1481 | 14 | -1 | -4.24 | <.001 | -0.04 |
| Self-prioritisation | 1436 | 13 | Admin-prioritisation | 1481 | 14 | 0 | 1.27 | .202 | 0.01 |
| *Asian* | | | | | | | | | |
| Total response | 1697 | 16 | Self-prioritisation | 1583 | 15 | -1 | -3.01 | .003 | -0.03 |
| Total response | 1697 | 16 | Admin-prioritisation | 1638 | 15 | -1 | -1.56 | .119 | -0.02 |
| Self-prioritisation | 1583 | 15 | Admin-prioritisation | 1638 | 15 | 1 | 1.50 | .135 | 0.01 |
| *Other* | | | | | | | | | |
| Total response | 554 | 5 | Self-prioritisation | 443 | 4 | -1 | -4.84 | <.001 | -0.05 |
| Total response | 554 | 5 | Admin-prioritisation | 455 | 4 | -1 | -4.32 | <.001 | -0.04 |
| Self-prioritisation | 443 | 4 | Admin-prioritisation | 455 | 4 | 0 | 0.58 | .560 | 0.01 |

[a]Cohen's (1988) *h* effect size.

APPENDIX B

# Supplementary Tables for Study 2

**Supplementary Table B.1**

*Demographic Characteristics by Ethnic Combination—Children[a] (N = 1,860)*

| | Māori/ European | | Pacific/ European | | Asian/ European | | MELAA/ European | | Māori/ Pacific | | Asian/ Other | | Māori/Pacific/ European | | Other combinations | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| *n* | 797 | | 202 | | 162 | | 63 | | 133 | | 82 | | 190 | | 231 | |
| Discrepancy between administrative-prioritisation and self-prioritisation, $\chi^2(14) = 114.09$, $p < .001$ | | | | | | | | | | | | | | | | |
| Non-discrepant | 167 | 21 | 66 | 33 | 10 | 6 | <10 | 6 | 53 | 40 | 30 | 37 | 57 | 30 | 77 | 33 |
| Discrepant—different main ethnic group | 426 | 53 | 90 | 45 | 104 | 64 | 46 | 73 | 36 | 27 | 31 | 38 | 77 | 41 | 90 | 39 |
| Discrepant—no main ethnic group | 204 | 26 | 46 | 23 | 48 | 30 | S | 21 | 44 | 33 | 21 | 26 | 56 | 29 | 64 | 28 |
| Sex, $\chi^2(7) = 3.93$, $p = .788$ | | | | | | | | | | | | | | | | |
| Female | 374 | 47 | 98 | 49 | 78 | 48 | 32 | 51 | 65 | 49 | 34 | 41 | 99 | 52 | 117 | 51 |
| Male | 423 | 53 | 104 | 51 | 84 | 52 | 31 | 49 | 68 | 51 | 48 | 59 | 91 | 48 | 114 | 49 |
| Mother's highest educational attainment, $\chi^2(14) = 133.52$, $p < .001$ | | | | | | | | | | | | | | | | |
| Secondary school qualification or below | 301 | 38 | 66 | 33 | 31 | 19 | 19 | 30 | 68 | 51 | 16 | 20 | 86 | 45 | 86 | 37 |
| Certificate or diploma | 250 | 31 | 81 | 40 | 41 | 25 | 11 | 17 | 52 | 39 | 24 | 29 | 71 | 37 | 83 | 36 |
| Bachelor's or above | 246 | 31 | 55 | 27 | 90 | 56 | 33 | 52 | 13 | 10 | 42 | 51 | 33 | 17 | 62 | 27 |
| Urbanicity, $\chi^2(14) = 133.00$, $p < .001$ | | | | | | | | | | | | | | | | |
| Main urban | 561 | 70 | 178 | 88 | 149 | 92 | 54 | 86 | 125 | 94 | 79 | 96 | 165 | 87 | 209 | 90 |
| Non-main urban | 129 | 16 | S | 8 | <10 | 4 | <10 | 6 | <10 | 5 | <10 | 2 | S | 9 | S | 7 |
| Rural | 107 | 13 | <10 | 3 | <10 | 4 | <10 | 8 | <10 | 2 | <10 | 1 | <10 | 4 | <10 | 3 |
| Socioeconomic deprivation, $\chi^2(14) = 191.87$, $p < .001$ | | | | | | | | | | | | | | | | |
| Low dep. | 199 | 25 | 48 | 24 | 64 | 40 | S | 40 | <10 | 5 | 23 | 28 | 16 | 8 | 38 | 16 |
| Medium dep. | 296 | 37 | 75 | 37 | 67 | 41 | 29 | 46 | S | 15 | 32 | 39 | 69 | 36 | 71 | 31 |
| High dep. | 302 | 38 | 79 | 39 | 31 | 19 | <10 | 14 | 107 | 80 | 27 | 33 | 105 | 55 | 122 | 53 |
| Area ethnic similarity (%),[b] $F(7, 1852) = 27.86$, $p < .001$ | | | | | | | | | | | | | | | | |
| Mean (*SD*) | 16.86 (10.91) | | 18.89 (19.59) | | 18.62 (13.92) | | 1.52 (0.86) | | 19.69 (9.06) | | 31.23 (16.13) | | 18.10 (11.43) | | 20.66 (17.40) | |
| Skewness | 1.28 | | 1.37 | | 0.91 | | 0.45 | | 1.06 | | 0.07 | | 1.84 | | 1.26 | |
| Kurtosis | 1.82 | | 0.88 | | 0.07 | | 0.53 | | 1.64 | | -0.88 | | 4.89 | | 0.98 | |

*Note*. MELAA = Middle Eastern, Latin American, and African. Cell counts less than 10 are suppressed as "<10". Secondary suppression (S) was applied to the next smallest cell so that the suppressed cell cannot be recalculated. Chi-square tests of independence were used to test for ethnic differences in categorical variables. A one-way analysis of variance (ANOVA) was used to test for ethnic differences in the continuous variable (i.e., area ethnic similarity).

[a]Child responses were collected by proxy from the child's mother.

[b]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; unstandardised.

**Supplementary Table B.2**

*Demographic Characteristics by Ethnic Combination—Adolescents (N = 2,413)*

| | Māori/ European | | Pacific/ European | | Asian/ European | | MELAA/ European | | Māori/ Pacific | | Pacific/ Asian | | Māori/Pacific/ European | | Other combinations | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| *n* | 1071 | | 418 | | 237 | | 88 | | 78 | | 103 | | 138 | | 279 | |
| Discrepancy between administrative-prioritisation and self-prioritisation, $\chi^2(14) = 157.61$, $p < .001$ | | | | | | | | | | | | | | | | |
| Non-discrepant | 379 | 35 | 241 | 58 | 100 | 42 | 28 | 32 | 47 | 60 | S | 37 | S | 48 | 178 | 64 |
| Discrepant—different main ethnic group | 664 | 62 | 167 | 40 | 117 | 49 | S | 59 | S | 37 | 59 | 58 | 68 | 49 | 87 | 31 |
| Discrepant—no main ethnic group | 28 | 3 | 10 | 2 | 20 | 8 | <10 | 9 | <10 | 3 | <10 | 5 | <10 | 3 | 14 | 5 |
| Sex, $\chi^2(7) = 12.99$, $p = .072$ | | | | | | | | | | | | | | | | |
| Female | 560 | 52 | 222 | 53 | 123 | 52 | 39 | 44 | 45 | 58 | 60 | 59 | 74 | 53 | 172 | 61 |
| Male | 512 | 48 | 196 | 47 | 114 | 48 | 49 | 56 | 33 | 42 | 42 | 41 | 64 | 47 | 108 | 39 |
| Age group (years), $\chi^2(7) = 12.80$, $p = .077$ | | | | | | | | | | | | | | | | |
| <16 | 747 | 70 | 289 | 69 | 145 | 61 | 61 | 69 | 61 | 79 | 66 | 64 | 99 | 72 | 186 | 67 |
| ≥16 | 324 | 30 | 129 | 31 | 92 | 39 | 28 | 31 | 17 | 21 | 37 | 36 | 39 | 28 | 93 | 33 |
| Birthplace, $\chi^2(7) = 591.63$, $p < .001$ | | | | | | | | | | | | | | | | |
| Aotearoa New Zealand | 1037 | 97 | 330 | 79 | 163 | 69 | 25 | 28 | S | 99 | 42 | 41 | S | 93 | 165 | 59 |
| Overseas | 35 | 3 | 88 | 21 | 74 | 31 | 63 | 72 | <10 | 1 | 60 | 59 | <10 | 7 | 114 | 41 |
| Urbanicity, $\chi^2(14) = 202.26$, $p < .001$ | | | | | | | | | | | | | | | | |
| Main urban | 695 | 65 | 378 | 90 | 212 | 89 | 74 | 84 | 63 | 81 | 101 | 98 | 104 | 75 | 240 | 86 |
| Non-main urban | 173 | 16 | 21 | 5 | 13 | 5 | <10 | 8 | S | 14 | <10 | <1 | 13 | 10 | 24 | 9 |
| Rural | 203 | 19 | 19 | 5 | 13 | 5 | <10 | 8 | <10 | 5 | <10 | 2 | 21 | 15 | 16 | 6 |
| Socioeconomic deprivation, $\chi^2(14) = 220.49$, $p < .001$ | | | | | | | | | | | | | | | | |
| Low dep. | 218 | 20 | 35 | 8 | 82 | 35 | 31 | 35 | <10 | 6 | <10 | 9 | 23 | 17 | 34 | 12 |
| Medium dep. | 416 | 39 | 113 | 27 | 101 | 42 | 44 | 50 | S | 23 | S | 33 | 42 | 31 | 90 | 32 |
| High dep. | 437 | 41 | 270 | 65 | 55 | 23 | 13 | 14 | 55 | 70 | 60 | 59 | 72 | 52 | 156 | 56 |
| Area ethnic similarity (%),[a] $F(7, 2405) = 40.97$, $p < .001$ | | | | | | | | | | | | | | | | |
| Mean (*SD*) | 19.92 (14.45) | | 30.90 (24.82) | | 19.46 (15.63) | | 1.42 (1.21) | | 19.63 (11.89) | | 28.86 (22.51) | | 19.80 (14.06) | | 26.62 (20.99) | |
| Skewness | 1.54 | | 0.28 | | 0.72 | | 1.33 | | 1.81 | | 0.48 | | 1.63 | | 0.77 | |
| Kurtosis | 2.41 | | -1.38 | | -0.51 | | 1.56 | | 3.42 | | -1.08 | | 3.11 | | -0.39 | |

*Note*. MELAA = Middle Eastern, Latin American, and African. Cell counts less than 10 are suppressed as "<10". Secondary suppression (S) was applied to the next smallest cell so that the suppressed cell cannot be recalculated. Chi-square tests of independence were used to test for ethnic differences in categorical variables. A one-way analysis of variance (ANOVA) was used to test for ethnic differences in the continuous variable (i.e., area ethnic similarity).

[a]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; unstandardised.

**Supplementary Table B.3**

*Demographic Characteristics by Ethnic Combination—Adults (N = 1,056)*

| | Māori/ European | | Pacific/ European | | Māori/ Pacific | | Māori/Pacific/ European | | Other combinations | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % | *n* | % |
| *n* | 627 | | 130 | | 72 | | 63 | | 164 | |
| Discrepancy between administrative-prioritisation and self-prioritisation, $\chi^2(8) = 7.61$, $p = .472$ | | | | | | | | | | |
| Non-discrepant | 246 | 39 | 50 | 38 | 35 | 49 | 23 | 37 | 62 | 38 |
| Discrepant—different main ethnic group | 246 | 39 | 44 | 34 | 23 | 32 | 27 | 43 | 71 | 43 |
| Discrepant—no main ethnic group | 135 | 22 | 36 | 28 | 14 | 19 | 13 | 21 | 31 | 19 |
| Age group (years), $\chi^2(4) = 38.88$, $p < .001$ | | | | | | | | | | |
| <30 | 358 | 57 | 81 | 62 | 59 | 82 | 50 | 79 | 76 | 46 |
| ≥30 | 269 | 43 | 49 | 38 | 13 | 18 | 13 | 21 | 88 | 54 |
| Birthplace, $\chi^2(4) = 266.29$, $p < .001$ | | | | | | | | | | |
| Aotearoa New Zealand | 609 | 97 | 107 | 82 | S | 100 | S | 95 | 87 | 53 |
| Overseas | 18 | 3 | 23 | 18 | <10 | <1 | <10 | 5 | 77 | 47 |
| Highest educational attainment, $\chi^2(8) = 52.28$, $p < .001$ | | | | | | | | | | |
| Secondary school qualification or below | 278 | 44 | 47 | 36 | 39 | 54 | 28 | 44 | 40 | 24 |
| Certificate or diploma | 204 | 33 | 57 | 44 | S | 33 | S | 41 | 54 | 33 |
| Bachelor's or above | 145 | 23 | 26 | 20 | <10 | 13 | <10 | 14 | 70 | 43 |
| Urbanicity, $\chi^2(8) = 45.23$, $p < .001$ | | | | | | | | | | |
| Main urban | 471 | 75 | 119 | 92 | 68 | 94 | 55 | 87 | 149 | 91 |
| Non-main urban | 98 | 16 | <10 | 5 | <10 | 3 | <10 | 10 | <10 | 5 |
| Rural | 58 | 9 | <10 | 3 | <10 | 3 | <10 | 3 | <10 | 4 |
| Socioeconomic deprivation, $\chi^2(8) = 53.18$, $p < .001$ | | | | | | | | | | |
| Low dep. | 117 | 19 | 17 | 13 | <10 | 3 | <10 | 11 | 43 | 26 |
| Medium dep. | 220 | 35 | 48 | 37 | S | 14 | S | 27 | 57 | 35 |
| High dep. | 290 | 46 | 65 | 50 | 60 | 83 | 39 | 62 | 64 | 39 |
| Area ethnic similarity (%),[a] $F(4, 1051) = 4.63$, $p = .001$ | | | | | | | | | | |
| Mean (*SD*) | 18.47 (10.94) | | 24.03 (22.05) | | 18.50 (8.88) | | 18.34 (10.93) | | 19.27 (16.58) | |
| Skewness | 1.06 | | 0.85 | | 1.45 | | 1.32 | | 1.04 | |
| Kurtosis | 1.28 | | -0.46 | | 3.57 | | 1.96 | | 0.61 | |

*Note*. MELAA = Middle Eastern, Latin American, and African. Cell counts less than 10 are suppressed as "<10". Secondary suppression (S) was applied to the next smallest cell so that the suppressed cell cannot be recalculated. Chi-square tests of independence were used to test for ethnic differences in categorical variables. A one-way analysis of variance (ANOVA) was used to test for ethnic differences in the continuous variable (i.e., area ethnic similarity).

[a]Percentage of neighbourhood area unit with the same administratively-prioritised ethnicity; unstandardised.

APPENDIX C

# Supplementary Tables for Study 3

**Supplementary Table C.1**

*Demographic Characteristics and Mental Health Outcomes by Ethnicity*

| | | Demographic characteristic | | | | | | | Mental health outcome[a] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age (years) | Sex | | Urbanicity | | | NZDep decile | Total difficulties score[b] | Self-harm | Suicide attempt |
| | | | Male | Female | Main urban | Minor urban | Rural | | | | |
| Ethnic group(s) | *n* | *M (SD)* | *n* (%) | *n* (%) | *n* (%) | *n* (%) | *n* (%) | *M (SD)* | *M (SD)* | *n* (%) | *n* (%) |
| Overall | 8275 | 14.89 (1.47) | 3752 (45) | 4523 (55) | 6158 (74) | 916 (11) | 1201 (15) | 5.63 (2.97) | 0.00 (1.00) | 1953 (24) | 364 (4) |
| **Sole/combination grouping** | | | | | | | | | | | |
| European | 3907 | 14.94 (1.46) | 1730 (44) | 2177 (56) | 2582 (66) | 525 (13) | 800 (20) | 4.47 (2.62) | -0.10 (1.01) | 853 (22) | 103 (3) |
| Māori | 288 | 14.65 (1.42) | 146 (51) | 142 (49) | 178 (62) | 60 (21) | 50 (17) | 7.83 (2.37) | 0.11 (0.93) | 66 (24) | 16 (6) |
| Pacific | 538 | 14.84 (1.46) | 224 (42) | 314 (58) | 517 (96) | S (4) | <10 (0) | 8.95 (1.67) | 0.06 (0.93) | 131 (25) | 43 (8) |
| Asian | 738 | 15.26 (1.53) | 367 (50) | 371 (50) | 700 (95) | 19 (3) | 18 (2) | 5.41 (2.71) | -0.08 (0.85) | 128 (17) | 21 (3) |
| Other | 159 | 15.24 (1.64) | 74 (47) | 85 (53) | 142 (90) | S (6) | <10 (4) | 6.27 (2.67) | 0.05 (0.96) | 31 (20) | <10 (4) |
| Māori/European | 967 | 14.75 (1.42) | 457 (47) | 510 (53) | 625 (65) | 163 (17) | 179 (18) | 6.27 (2.81) | 0.10 (0.98) | 262 (27) | 49 (5) |
| Pacific/European | 384 | 14.70 (1.50) | 179 (47) | 205 (53) | 350 (91) | 16 (4) | 17 (4) | 7.86 (2.52) | 0.12 (1.00) | 98 (26) | 29 (8) |
| Asian/European | 224 | 15.00 (1.70) | 106 (47) | 118 (53) | 199 (89) | 13 (6) | 13 (6) | 4.90 (2.63) | 0.04 (1.02) | 51 (23) | 12 (6) |
| Māori/Pacific | 78 | 14.59 (1.32) | 33 (42) | 45 (58) | 63 (81) | S (14) | <10 (5) | 7.95 (2.30) | 0.31 (1.04) | 21 (26) | <10 (10) |
| Māori/Pacific/European | 125 | 14.47 (1.36) | 57 (46) | 68 (54) | 95 (76) | 13 (11) | 17 (14) | 6.92 (2.67) | 0.30 (1.04) | 43 (37) | 11 (9) |
| 2 groups NEI | 559 | 14.74 (1.43) | 236 (42) | 323 (58) | 467 (83) | 37 (7) | 56 (10) | 5.97 (2.94) | 0.08 (0.99) | 161 (29) | 38 (7) |
| ≥3 groups NEI | 309 | 14.70 (1.41) | 144 (47) | 165 (53) | 240 (78) | 28 (9) | 41 (13) | 6.80 (2.89) | 0.35 (1.09) | 108 (36) | 28 (10) |
| **Total response grouping** | | | | | | | | | | | |
| European | 6209 | 14.86 (1.46) | 2806 (45) | 3403 (55) | 4312 (69) | 783 (13) | 1113 (18) | 5.14 (2.86) | -0.01 (1.02) | 1500 (24) | 246 (4) |
| Māori | 1669 | 14.70 (1.41) | 792 (47) | 877 (53) | 1107 (66) | 274 (16) | 288 (17) | 6.73 (2.78) | 0.16 (1.00) | 469 (29) | 105 (6) |
| Pacific | 1422 | 14.75 (1.44) | 610 (43) | 812 (57) | 1299 (91) | 71 (5) | 52 (4) | 8.16 (2.33) | 0.12 (0.98) | 379 (27) | 119 (9) |
| Asian | 1256 | 15.12 (1.56) | 592 (47) | 664 (53) | 1169 (93) | 46 (4) | 41 (3) | 5.68 (2.79) | -0.03 (0.93) | 272 (22) | 58 (5) |
| Other | 839 | 14.77 (1.48) | 378 (45) | 461 (55) | 673 (80) | 66 (8) | 100 (12) | 6.07 (2.93) | 0.19 (1.01) | 244 (30) | 57 (7) |
| **Administrative-prioritisation** | | | | | | | | | | | |
| European | 3907 | 14.94 (1.46) | 1730 (44) | 2177 (56) | 2582 (66) | 525 (13) | 800 (20) | 4.47 (2.62) | -0.10 (1.01) | 853 (22) | 103 (3) |
| Māori | 1669 | 14.70 (1.41) | 792 (47) | 877 (53) | 1107 (66) | 274 (16) | 288 (17) | 6.73 (2.78) | 0.16 (1.00) | 469 (29) | 105 (6) |
| Pacific | 1180 | 14.80 (1.46) | 503 (43) | 677 (57) | 1110 (94) | 45 (4) | 25 (2) | 8.33 (2.24) | 0.07 (0.96) | 301 (26) | 95 (8) |
| Asian | 1036 | 15.18 (1.58) | 509 (49) | 527 (51) | 968 (93) | 34 (3) | 34 (3) | 5.35 (2.71) | -0.04 (0.91) | 203 (20) | 37 (4) |
| Other | 484 | 14.80 (1.50) | 219 (45) | 265 (55) | 392 (81) | 37 (8) | 54 (11) | 5.31 (2.78) | 0.11 (0.98) | 128 (27) | 25 (5) |
| **Self-prioritisation** | | | | | | | | | | | |
| European | 5181 | 14.88 (1.46) | 2323 (45) | 2859 (55) | 3508 (68) | 676 (13) | 998 (19) | 4.76 (2.73) | -0.04 (1.02) | 1215 (24) | 169 (3) |
| Māori | 832 | 14.65 (1.38) | 401 (48) | 430 (52) | 528 (64) | 154 (18) | 150 (18) | 7.44 (2.47) | 0.14 (0.97) | 231 (29) | 57 (7) |
| Pacific | 967 | 14.81 (1.44) | 400 (41) | 568 (59) | 929 (96) | S (3) | <10 (1) | 8.76 (1.89) | 0.11 (0.94) | 250 (26) | 84 (9) |
| Asian | 949 | 15.23 (1.57) | 474 (50) | 475 (50) | 898 (95) | 26 (3) | 25 (3) | 5.52 (2.71) | -0.06 (0.89) | 177 (19) | 33 (4) |
| Other | 239 | 15.11 (1.63) | 117 (49) | 122 (51) | 212 (89) | 15 (6) | 12 (5) | 6.00 (2.73) | 0.03 (0.93) | 50 (22) | 15 (7) |
| Can't choose one ethnic group | 107 | 14.60 (1.53) | 37 (35) | 70 (65) | 84 (78) | 13 (12) | 10 (9) | 5.61 (3.08) | 0.12 (1.03) | 30 (29) | <10 (6) |

*Note*. Cell counts less than 10 are suppressed as "<10". Secondary suppression (S) was applied to the next smallest cell so that the suppressed cell cannot be recalculated. NEI = not elsewhere included.

[a]Complete case analysis used (total difficulties $N$ = 7,990 [97%]; self-harm $N$ = 8,170 [99%]; suicide attempt $N$ = 8,119 [98%]).

[b]Total difficulties score was standardised ($M$ = 0, $SD$ = 1).

**Supplementary Table C.2**

*Multiple Linear Regression Coefficients for Total Difficulties Score[a] Using Non-Mutually Exclusive Ethnic Classification Methods*

| | European model | | | Māori model | | | Pacific model | | | Asian model | | | Other ethnicities model | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | *B* | *SE* | *p* | *B* | *SE* | *p* | *B* | *SE* | *p* | *B* | *SE* | *p* | *B* | *SE* | *p* |
| **Original total response models** | | | | | | | | | | | | | | | |
| Intercept | -0.33 | 0.04 | <.001 | -0.32 | 0.04 | <.001 | -0.30 | 0.04 | <.001 | -0.29 | 0.04 | <.001 | -0.31 | 0.04 | <.001 |
| Female (ref. male) | 0.19 | 0.02 | <.001 | 0.19 | 0.02 | <.001 | 0.19 | 0.02 | <.001 | 0.19 | 0.02 | <.001 | 0.19 | 0.02 | <.001 |
| Age 14 (ref. age ≤13 years) | 0.05 | 0.03 | .152 | 0.05 | 0.03 | .152 | 0.05 | 0.03 | .152 | 0.05 | 0.03 | .152 | 0.05 | 0.03 | .163 |
| Age 15 (ref. age ≤13 years) | 0.12 | 0.03 | .001 | 0.12 | 0.03 | <.001 | 0.12 | 0.03 | .001 | 0.12 | 0.03 | .001 | 0.12 | 0.03 | <.001 |
| Age 16 (ref. age ≤13 years) | 0.13 | 0.03 | <.001 | 0.13 | 0.03 | <.001 | 0.13 | 0.03 | <.001 | 0.13 | 0.03 | <.001 | 0.13 | 0.03 | <.001 |
| Age ≥17 (ref. age ≤13 years) | 0.02 | 0.04 | .501 | 0.03 | 0.04 | .348 | 0.02 | 0.04 | .526 | 0.02 | 0.04 | .506 | 0.03 | 0.04 | .485 |
| Minor urban (ref. main urban) | 0.01 | 0.04 | .814 | 0.00 | 0.04 | .985 | 0.03 | 0.04 | .486 | 0.01 | 0.04 | .746 | 0.02 | 0.04 | .485 |
| Rural (ref. main urban) | -0.05 | 0.03 | .090 | -0.06 | 0.03 | .057 | -0.04 | 0.03 | .193 | -0.05 | 0.03 | .103 | -0.04 | 0.03 | .172 |
| NZDep3–4 (ref. NZDep1–2) | 0.07 | 0.04 | .051 | 0.06 | 0.03 | .073 | 0.06 | 0.03 | .074 | 0.07 | 0.04 | .057 | 0.06 | 0.03 | .081 |
| NZDep5–6 (ref. NZDep1–2) | 0.12 | 0.04 | .001 | 0.10 | 0.04 | .003 | 0.11 | 0.04 | .002 | 0.12 | 0.04 | .001 | 0.11 | 0.04 | .001 |
| NZDep7–8 (ref. NZDep1–2) | 0.18 | 0.04 | <.001 | 0.15 | 0.04 | <.001 | 0.17 | 0.04 | <.001 | 0.18 | 0.04 | <.001 | 0.17 | 0.04 | <.001 |
| NZDep9–10 (ref. NZDep1–2) | 0.28 | 0.04 | <.001 | 0.23 | 0.03 | <.001 | 0.24 | 0.04 | <.001 | 0.26 | 0.03 | <.001 | 0.26 | 0.03 | <.001 |
| European (ref. non-European) | 0.05 | 0.03 | .095 | | - | | | - | | | - | | | - | |
| Māori (ref. non-Māori) | | - | | 0.18 | 0.03 | <.001 | | - | | | - | | | - | |
| Pacific (ref. non-Pacific) | | - | | | - | | 0.05 | 0.03 | .129 | | - | | | - | |
| Asian (ref. non-Asian) | | - | | | - | | | - | | -0.04 | 0.03 | .241 | | - | |
| Other (ref. non-Other) | | - | | | - | | | - | | | - | | 0.20 | 0.04 | <.001 |
| **Modified total response models** | | | | | | | | | | | | | | | |
| Intercept | | | | -0.38 | 0.04 | <.001 | -0.31 | 0.04 | <.001 | -0.35 | 0.04 | <.001 | -0.37 | 0.04 | <.001 |
| Female (ref. male) | | | | 0.22 | 0.03 | <.001 | 0.16 | 0.03 | <.001 | 0.19 | 0.03 | <.001 | 0.20 | 0.03 | <.001 |
| Age 14 (ref. age ≤13 years) | | | | 0.05 | 0.04 | .181 | 0.01 | 0.04 | .784 | 0.03 | 0.04 | .533 | 0.05 | 0.04 | .296 |
| Age 15 (ref. age ≤13 years) | | | | 0.15 | 0.04 | <.001 | 0.12 | 0.04 | .007 | 0.14 | 0.04 | .001 | 0.16 | 0.05 | .001 |
| Age 16 (ref. age ≤13 years) | | | | 0.15 | 0.04 | <.001 | 0.15 | 0.04 | .001 | 0.17 | 0.04 | <.001 | 0.18 | 0.05 | <.001 |
| Age ≥17 (ref. age ≤13 years) | | | | 0.00 | 0.04 | .974 | -0.02 | 0.05 | .710 | 0.08 | 0.04 | .066 | 0.03 | 0.05 | .594 |
| Minor urban (ref. main urban) | | | | -0.01 | 0.04 | .756 | 0.03 | 0.05 | .573 | -0.03 | 0.05 | .539 | 0.01 | 0.05 | .821 |
| Rural (ref. main urban) | | -[b] | | -0.06 | 0.04 | .089 | -0.05 | 0.04 | .210 | -0.08 | 0.04 | .043 | -0.05 | 0.04 | .202 |
| NZDep3–4 (ref. NZDep1–2) | | | | 0.07 | 0.04 | .107 | 0.05 | 0.04 | .254 | 0.05 | 0.04 | .245 | 0.06 | 0.04 | .127 |
| NZDep5–6 (ref. NZDep1–2) | | | | 0.13 | 0.04 | .002 | 0.10 | 0.04 | .026 | 0.11 | 0.04 | .005 | 0.11 | 0.04 | .011 |
| NZDep7–8 (ref. NZDep1–2) | | | | 0.19 | 0.04 | <.001 | 0.17 | 0.05 | <.001 | 0.13 | 0.04 | .004 | 0.16 | 0.05 | .001 |
| NZDep9–10 (ref. NZDep1–2) | | | | 0.25 | 0.05 | <.001 | 0.20 | 0.05 | <.001 | 0.28 | 0.05 | <.001 | 0.24 | 0.05 | <.001 |
| Māori (ref. sole European) | | | | 0.20 | 0.03 | <.001 | | - | | | - | | | - | |
| Pacific (ref. sole European) | | | | | - | | 0.12 | 0.04 | .002 | | - | | | - | |
| Asian (ref. sole European) | | | | | - | | | - | | 0.02 | 0.03 | .650 | | - | |
| Other (ref. sole European) | | | | | - | | | - | | | - | | 0.25 | 0.04 | <.001 |

[a]Total difficulties score was standardised ($M = 0$, $SD = 1$).

[b]Model not applicable as the reference group for modified total response is sole European.

**Supplementary Table C.3**

*Multiple Linear Regression Coefficients for Total Difficulties Score[a] Using Mutually Exclusive Ethnic Classification Methods*

| | Sole/combination model | | | Administrative-prioritisation model | | | Self-prioritisation model | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | *B* | *SE* | *p* | *B* | *SE* | *p* | *B* | *SE* | *p* |
| Intercept | -0.34 | 0.04 | <.001 | -0.34 | 0.04 | <.001 | -0.30 | 0.04 | <.001 |
| Female (ref. male) | 0.19 | 0.02 | <.001 | 0.19 | 0.02 | <.001 | 0.19 | 0.02 | <.001 |
| Age 14 (ref. age ≤13 years) | 0.05 | 0.03 | .165 | 0.05 | 0.03 | .167 | 0.05 | 0.03 | .163 |
| Age 15 (ref. age ≤13 years) | 0.12 | 0.03 | <.001 | 0.12 | 0.03 | <.001 | 0.12 | 0.03 | .001 |
| Age 16 (ref. age ≤13 years) | 0.14 | 0.03 | <.001 | 0.14 | 0.03 | <.001 | 0.13 | 0.03 | <.001 |
| Age ≥17 (ref. age ≤13 years) | 0.04 | 0.04 | .265 | 0.03 | 0.04 | .336 | 0.03 | 0.04 | .426 |
| Minor urban (ref. main urban) | 0.03 | 0.04 | .472 | 0.02 | 0.04 | .655 | 0.01 | 0.04 | .809 |
| Rural (ref. main urban) | -0.04 | 0.03 | .177 | -0.05 | 0.03 | .138 | -0.06 | 0.03 | .088 |
| NZDep3–4 (ref. NZDep1–2) | 0.06 | 0.03 | .115 | 0.06 | 0.03 | .113 | 0.06 | 0.04 | .066 |
| NZDep5–6 (ref. NZDep1–2) | 0.10 | 0.04 | .007 | 0.10 | 0.04 | .007 | 0.11 | 0.04 | .002 |
| NZDep7–8 (ref. NZDep1–2) | 0.13 | 0.04 | <.001 | 0.13 | 0.04 | <.001 | 0.16 | 0.04 | <.001 |
| NZDep9–10 (ref. NZDep1–2) | 0.19 | 0.04 | <.001 | 0.20 | 0.04 | <.001 | 0.24 | 0.04 | <.001 |
| Māori (ref. European[b]) | 0.15 | 0.06 | .019 | 0.21 | 0.03 | <.001 | 0.12 | 0.04 | .003 |
| Pacific (ref. European[b]) | 0.05 | 0.05 | .384 | 0.07 | 0.04 | .066 | 0.02 | 0.04 | .601 |
| Asian (ref. European[b]) | -0.01 | 0.04 | .834 | 0.03 | 0.04 | .356 | -0.04 | 0.04 | .287 |
| Other (ref. European[b]) | 0.11 | 0.08 | .189 | 0.19 | 0.05 | <.001 | 0.04 | 0.07 | .556 |
| Māori/European (ref. European[b]) | 0.16 | 0.04 | <.001 | | - | | | - | |
| Pacific/European (ref. European[b]) | 0.14 | 0.06 | .011 | | - | | | - | |
| Asian/European (ref. European[b]) | 0.13 | 0.07 | .054 | | - | | | - | |
| Māori/Pacific (ref. European[b]) | 0.32 | 0.12 | .005 | | - | | | - | |
| Māori/Pacific/European (ref. European[b]) | 0.35 | 0.09 | <.001 | | - | | | - | |
| 2 groups NEI (ref. European[b]) | 0.14 | 0.05 | .003 | | - | | | - | |
| ≥3 groups NEI (ref. European[b]) | 0.40 | 0.06 | <.001 | | - | | | - | |
| Can't choose one ethnic group (ref. European[b]) | | - | | | - | | 0.12 | 0.10 | .231 |

*Note*. NEI = not elsewhere included.

[a]Total difficulties score was standardised ($M = 0$, $SD = 1$).

[b]For the sole/combination and administrative-prioritisation models, the referent is *sole* European; for the self-prioritisation model, the referent is *self-prioritised* European.

**Supplementary Table C.4**

*Binary Logistic Regression Coefficients for Self-Harm[a] Using Non-Mutually Exclusive Ethnic Classification Methods*

| Variable | European model | | | Māori model | | | Pacific model | | | Asian model | | | Other ethnicities model | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* |
| **Original total response models** | | | | | | | | | | | | | | | |
| Intercept | 0.13 | 0.11 | <.001 | 0.15 | 0.09 | <.001 | 0.16 | 0.09 | <.001 | 0.16 | 0.09 | <.001 | 0.15 | 0.09 | <.001 |
| Female (ref. male) | 1.90 | 0.05 | <.001 | 1.90 | 0.05 | <.001 | 1.89 | 0.05 | <.001 | 1.89 | 0.05 | <.001 | 1.89 | 0.05 | <.001 |
| Age 14 (ref. age ≤13 years) | 1.18 | 0.08 | .043 | 1.18 | 0.08 | .044 | 1.18 | 0.08 | .044 | 1.18 | 0.08 | .044 | 1.17 | 0.08 | .048 |
| Age 15 (ref. age ≤13 years) | 1.28 | 0.08 | .002 | 1.28 | 0.08 | .002 | 1.27 | 0.08 | .003 | 1.27 | 0.08 | .003 | 1.28 | 0.08 | .003 |
| Age 16 (ref. age ≤13 years) | 1.19 | 0.08 | .035 | 1.20 | 0.08 | .032 | 1.18 | 0.08 | .045 | 1.19 | 0.08 | .040 | 1.19 | 0.08 | .038 |
| Age ≥17 (ref. age ≤13 years) | 1.00 | 0.09 | .981 | 1.00 | 0.09 | .957 | 0.98 | 0.09 | .861 | 0.99 | 0.09 | .924 | 0.99 | 0.09 | .896 |
| Minor urban (ref. main urban) | 0.97 | 0.08 | .742 | 0.99 | 0.08 | .882 | 1.03 | 0.08 | .699 | 1.00 | 0.08 | .972 | 1.03 | 0.08 | .708 |
| Rural (ref. main urban) | 0.87 | 0.08 | .080 | 0.88 | 0.08 | .111 | 0.91 | 0.08 | .260 | 0.89 | 0.08 | .128 | 0.91 | 0.08 | .242 |
| NZDep3–4 (ref. NZDep1–2) | 1.18 | 0.09 | .051 | 1.16 | 0.09 | .090 | 1.16 | 0.09 | .092 | 1.17 | 0.09 | .069 | 1.15 | 0.09 | .098 |
| NZDep5–6 (ref. NZDep1–2) | 1.25 | 0.09 | .009 | 1.20 | 0.09 | .036 | 1.21 | 0.09 | .027 | 1.23 | 0.09 | .017 | 1.22 | 0.09 | .024 |
| NZDep7–8 (ref. NZDep1–2) | 1.48 | 0.09 | <.001 | 1.35 | 0.09 | <.001 | 1.39 | 0.09 | <.001 | 1.43 | 0.09 | <.001 | 1.39 | 0.09 | <.001 |
| NZDep9–10 (ref. NZDep1–2) | 1.57 | 0.08 | <.001 | 1.36 | 0.08 | <.001 | 1.38 | 0.09 | <.001 | 1.44 | 0.08 | <.001 | 1.42 | 0.08 | <.001 |
| European (ref. non-European) | 1.27 | 0.07 | <.001 | | - | | | - | | | - | | | - | |
| Māori (ref. non-Māori) | | - | | 1.33 | 0.06 | <.001 | | - | | | - | | | - | |
| Pacific (ref. non-Pacific) | | - | | | - | | 1.09 | 0.08 | .238 | | - | | | - | |
| Asian (ref. non-Asian) | | - | | | - | | | - | | 0.87 | 0.08 | .066 | | - | |
| Other (ref. non-Other) | | - | | | - | | | - | | | - | | 1.37 | 0.08 | <.001 |
| **Modified total response models** | | | | | | | | | | | | | | | |
| Intercept | | | | 0.14 | 0.11 | <.001 | 0.14 | 0.11 | <.001 | 0.14 | 0.11 | <.001 | 0.13 | 0.12 | <.001 |
| Female (ref. male) | | | | 1.99 | 0.07 | <.001 | 1.79 | 0.07 | <.001 | 1.83 | 0.07 | <.001 | 1.89 | 0.07 | <.001 |
| Age 14 (ref. age ≤13 years) | | | | 1.16 | 0.10 | .131 | 1.15 | 0.10 | .171 | 1.11 | 0.11 | .350 | 1.30 | 0.11 | .016 |
| Age 15 (ref. age ≤13 years) | | | | 1.43 | 0.10 | <.001 | 1.36 | 0.10 | .003 | 1.32 | 0.11 | .009 | 1.47 | 0.11 | <.001 |
| Age 16 (ref. age ≤13 years) | | | | 1.24 | 0.10 | .034 | 1.37 | 0.10 | .002 | 1.25 | 0.11 | .041 | 1.40 | 0.11 | .003 |
| Age ≥17 (ref. age ≤13 years) | | | | 1.03 | 0.11 | .772 | 1.08 | 0.11 | .512 | 1.16 | 0.11 | .183 | 1.18 | 0.12 | .169 |
| Minor urban (ref. main urban) | | | | 0.93 | 0.09 | .470 | 1.00 | 0.11 | .982 | 0.95 | 0.11 | .623 | 1.03 | 0.11 | .793 |
| Rural (ref. main urban) | | -[b] | | 0.81 | 0.09 | .017 | 0.88 | 0.10 | .210 | 0.88 | 0.10 | .205 | 0.92 | 0.09 | .391 |
| NZDep3–4 (ref. NZDep1–2) | | | | 1.19 | 0.10 | .086 | 1.22 | 0.11 | .063 | 1.19 | 0.10 | .085 | 1.20 | 0.10 | .071 |
| NZDep5–6 (ref. NZDep1–2) | | | | 1.30 | 0.10 | .009 | 1.33 | 0.11 | .008 | 1.34 | 0.10 | .004 | 1.24 | 0.11 | .040 |
| NZDep7–8 (ref. NZDep1–2) | | | | 1.55 | 0.10 | <.001 | 1.58 | 0.11 | <.001 | 1.56 | 0.11 | <.001 | 1.53 | 0.11 | <.001 |
| NZDep9–10 (ref. NZDep1–2) | | | | 1.48 | 0.11 | <.001 | 1.47 | 0.12 | .001 | 1.69 | 0.12 | <.001 | 1.51 | 0.12 | .001 |
| Māori (ref. sole European) | | | | 1.32 | 0.07 | <.001 | | - | | | - | | | - | |
| Pacific (ref. sole European) | | | | | - | | 1.14 | 0.09 | .165 | | - | | | - | |
| Asian (ref. sole European) | | | | | - | | | - | | 0.91 | 0.08 | .253 | | - | |
| Other (ref. sole European) | | | | | - | | | - | | | - | | 1.41 | 0.09 | <.001 |

[a]0 = no self-harm in the past 12 months; 1 = self-harm in the past 12 months.

[b]Model not applicable as the reference group for modified total response is sole European.

**Supplementary Table C.5**

*Binary Logistic Regression Coefficients for Self-Harm[a] Using Mutually Exclusive Ethnic Classification Methods*

| Variable | Sole/combination model | | | Administrative-prioritisation model | | | Self-prioritisation model | | |
|---|---|---|---|---|---|---|---|---|---|
| | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* |
| Intercept | 0.15 | 0.09 | <.001 | 0.15 | 0.09 | <.001 | 0.16 | 0.09 | <.001 |
| Female (ref. male) | 1.89 | 0.05 | <.001 | 1.90 | 0.05 | <.001 | 1.89 | 0.05 | <.001 |
| Age 14 (ref. age ≤13 years) | 1.17 | 0.08 | .050 | 1.17 | 0.08 | .049 | 1.17 | 0.08 | .045 |
| Age 15 (ref. age ≤13 years) | 1.30 | 0.08 | .001 | 1.28 | 0.08 | .002 | 1.28 | 0.08 | .002 |
| Age 16 (ref. age ≤13 years) | 1.22 | 0.08 | .017 | 1.21 | 0.08 | .023 | 1.21 | 0.08 | .026 |
| Age ≥17 (ref. age ≤13 years) | 1.04 | 0.09 | .669 | 1.02 | 0.09 | .858 | 1.02 | 0.09 | .856 |
| Minor urban (ref. main urban) | 1.00 | 0.09 | .988 | 0.99 | 0.09 | .892 | 0.96 | 0.09 | .615 |
| Rural (ref. main urban) | 0.88 | 0.08 | .100 | 0.88 | 0.08 | .098 | 0.85 | 0.08 | .047 |
| NZDep3–4 (ref. NZDep1–2) | 1.16 | 0.09 | .081 | 1.16 | 0.09 | .091 | 1.18 | 0.09 | .052 |
| NZDep5–6 (ref. NZDep1–2) | 1.21 | 0.09 | .032 | 1.20 | 0.09 | .041 | 1.24 | 0.09 | .016 |
| NZDep7–8 (ref. NZDep1–2) | 1.34 | 0.09 | .001 | 1.34 | 0.09 | .001 | 1.42 | 0.09 | <.001 |
| NZDep9–10 (ref. NZDep1–2) | 1.34 | 0.09 | .001 | 1.32 | 0.09 | .002 | 1.43 | 0.09 | <.001 |
| Māori (ref. European[b]) | 1.04 | 0.15 | .795 | 1.36 | 0.07 | <.001 | 1.19 | 0.09 | .056 |
| Pacific (ref. European[b]) | 0.99 | 0.12 | .967 | 1.09 | 0.09 | .321 | 0.94 | 0.10 | .539 |
| Asian (ref. European[b]) | 0.73 | 0.11 | .003 | 0.86 | 0.09 | .088 | 0.72 | 0.09 | <.001 |
| Other (ref. European[b]) | 0.86 | 0.21 | .459 | 1.27 | 0.11 | .032 | 0.84 | 0.17 | .301 |
| Māori/European (ref. European[b]) | 1.29 | 0.09 | .003 | | - | | | - | |
| Pacific/European (ref. European[b]) | 1.16 | 0.13 | .265 | | - | | | - | |
| Asian/European (ref. European[b]) | 1.08 | 0.17 | .652 | | - | | | - | |
| Māori/Pacific (ref. European[b]) | 1.10 | 0.27 | .715 | | - | | | - | |
| Māori/Pacific/European (ref. European[b]) | 1.92 | 0.20 | .001 | | - | | | - | |
| 2 groups NEI (ref. European[b]) | 1.36 | 0.10 | .003 | | - | | | - | |
| ≥3 groups NEI (ref. European[b]) | 1.89 | 0.13 | <.001 | | - | | | - | |
| Can't choose one ethnic group (ref. European[b]) | | - | | | - | | 1.18 | 0.22 | .450 |

*Note*. NEI = not elsewhere included.

[a]0 = no self-harm in the past 12 months; 1 = self-harm in the past 12 months.

[b]For the sole/combination and administrative-prioritisation models, the referent is *sole* European; for the self-prioritisation model, the referent is *self-prioritised* European.

**Supplementary Table C.6**

*Binary Logistic Regression Coefficients for Suicide Attempt[a] Using Non-Mutually Exclusive Ethnic Classification Methods*

| | European model | | | Māori model | | | Pacific model | | | Asian model | | | Other ethnicities model | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* |
| **Original total response models** | | | | | | | | | | | | | | | |
| Intercept | 0.01 | 0.24 | <.001 | 0.01 | 0.21 | <.001 | 0.01 | 0.21 | <.001 | 0.01 | 0.21 | <.001 | 0.01 | 0.21 | <.001 |
| Female (ref. male) | 2.87 | 0.13 | <.001 | 2.91 | 0.13 | <.001 | 2.81 | 0.13 | <.001 | 2.88 | 0.13 | <.001 | 2.87 | 0.13 | <.001 |
| Age 14 (ref. age ≤13 years) | 1.56 | 0.16 | .007 | 1.56 | 0.16 | .007 | 1.56 | 0.16 | .007 | 1.56 | 0.16 | .007 | 1.56 | 0.16 | .007 |
| Age 15 (ref. age ≤13 years) | 1.25 | 0.17 | .204 | 1.27 | 0.17 | .174 | 1.26 | 0.17 | .188 | 1.25 | 0.17 | .203 | 1.25 | 0.17 | .193 |
| Age 16 (ref. age ≤13 years) | 1.31 | 0.18 | .122 | 1.35 | 0.18 | .087 | 1.33 | 0.18 | .104 | 1.32 | 0.18 | .121 | 1.34 | 0.18 | .102 |
| Age ≥17 (ref. age ≤13 years) | 0.95 | 0.19 | .811 | 1.00 | 0.19 | .997 | 0.99 | 0.19 | .942 | 0.96 | 0.19 | .817 | 0.98 | 0.19 | .900 |
| Minor urban (ref. main urban) | 1.01 | 0.17 | .954 | 0.93 | 0.17 | .669 | 1.15 | 0.17 | .423 | 1.00 | 0.17 | .977 | 1.00 | 0.17 | .976 |
| Rural (ref. main urban) | 1.03 | 0.17 | .859 | 0.96 | 0.17 | .791 | 1.11 | 0.17 | .540 | 1.02 | 0.17 | .899 | 1.01 | 0.17 | .935 |
| NZDep3–4 (ref. NZDep1–2) | 1.02 | 0.21 | .919 | 1.03 | 0.21 | .897 | 1.00 | 0.21 | .992 | 1.03 | 0.21 | .893 | 1.02 | 0.21 | .911 |
| NZDep5–6 (ref. NZDep1–2) | 1.40 | 0.19 | .084 | 1.39 | 0.19 | .094 | 1.33 | 0.20 | .148 | 1.42 | 0.19 | .073 | 1.41 | 0.19 | .074 |
| NZDep7–8 (ref. NZDep1–2) | 1.38 | 0.20 | .102 | 1.33 | 0.20 | .149 | 1.25 | 0.20 | .256 | 1.41 | 0.20 | .080 | 1.39 | 0.20 | .094 |
| NZDep9–10 (ref. NZDep1–2) | 2.55 | 0.18 | <.001 | 2.50 | 0.17 | <.001 | 1.98 | 0.18 | <.001 | 2.69 | 0.17 | <.001 | 2.63 | 0.17 | <.001 |
| European (ref. non-European) | 0.87 | 0.13 | .256 | | - | | | - | | | - | | | - | |
| Māori (ref. non-Māori) | | - | | 1.49 | 0.12 | .001 | | - | | | - | | | - | |
| Pacific (ref. non-Pacific) | | - | | | - | | 1.89 | 0.14 | <.001 | | - | | | - | |
| Asian (ref. non-Asian) | | - | | | - | | | - | | 1.11 | 0.15 | .490 | | - | |
| Other (ref. non-Other) | | - | | | - | | | - | | | - | | 1.62 | 0.15 | .002 |
| **Modified total response models** | | | | | | | | | | | | | | | |
| Intercept | | | | 0.01 | 0.27 | <.001 | 0.01 | 0.29 | <.001 | 0.01 | 0.31 | <.001 | 0.01 | 0.31 | <.001 |
| Female (ref. male) | | | | 2.87 | 0.17 | <.001 | 3.23 | 0.17 | <.001 | 2.59 | 0.19 | <.001 | 3.08 | 0.20 | <.001 |
| Age 14 (ref. age ≤13 years) | | | | 1.33 | 0.21 | .172 | 1.77 | 0.21 | .008 | 1.54 | 0.27 | .104 | 1.35 | 0.26 | .257 |
| Age 15 (ref. age ≤13 years) | | | | 1.12 | 0.22 | .616 | 1.29 | 0.23 | .264 | 1.39 | 0.27 | .222 | 1.31 | 0.27 | .311 |
| Age 16 (ref. age ≤13 years) | | | | 1.28 | 0.23 | .283 | 1.66 | 0.23 | .025 | 1.51 | 0.27 | .129 | 1.97 | 0.26 | .008 |
| Age ≥17 (ref. age ≤13 years) | | | | 0.95 | 0.25 | .849 | 0.96 | 0.26 | .876 | 1.26 | 0.28 | .406 | 1.10 | 0.30 | .742 |
| Minor urban (ref. main urban) | | | | 1.02 | 0.20 | .937 | 1.25 | 0.23 | .331 | 0.80 | 0.29 | .441 | 1.03 | 0.26 | .920 |
| Rural (ref. main urban) | | -[b] | | 1.11 | 0.19 | .565 | 1.35 | 0.23 | .184 | 1.48 | 0.22 | .079 | 1.45 | 0.21 | .079 |
| NZDep3–4 (ref. NZDep1–2) | | | | 1.05 | 0.25 | .846 | 0.79 | 0.29 | .413 | 0.91 | 0.27 | .732 | 0.78 | 0.27 | .351 |
| NZDep5–6 (ref. NZDep1–2) | | | | 1.43 | 0.23 | .128 | 1.54 | 0.25 | .088 | 1.79 | 0.24 | .016 | 1.42 | 0.24 | .152 |
| NZDep7–8 (ref. NZDep1–2) | | | | 1.34 | 0.24 | .225 | 1.25 | 0.27 | .410 | 1.30 | 0.28 | .339 | 1.19 | 0.27 | .515 |
| NZDep9–10 (ref. NZDep1–2) | | | | 1.59 | 0.24 | .057 | 1.83 | 0.26 | .018 | 2.66 | 0.26 | <.001 | 2.05 | 0.26 | .006 |
| Māori (ref. sole European) | | | | 2.31 | 0.16 | <.001 | | - | | | - | | | - | |
| Pacific (ref. sole European) | | | | | - | | 2.74 | 0.18 | <.001 | | - | | | - | |
| Asian (ref. sole European) | | | | | - | | | - | | 1.69 | 0.19 | .005 | | - | |
| Other (ref. sole European) | | | | | - | | | - | | | - | | 2.52 | 0.19 | <.001 |

[a]0 = no suicide attempt in the past 12 months; 1 = suicide attempt in the past 12 months.

[b]Model not applicable as the reference group for modified total response is sole European.

**Supplementary Table C.7**

*Binary Logistic Regression Coefficients for Suicide Attempt[a] Using Mutually Exclusive Ethnic Classification Methods*

| Variable | Sole/combination model | | | Administrative-prioritisation model | | | Self-prioritisation model | | |
|---|---|---|---|---|---|---|---|---|---|
| | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* | *Exp(B)* | *SE* | *p* |
| Intercept | 0.01 | 0.22 | <.001 | 0.01 | 0.22 | <.001 | 0.01 | 0.21 | <.001 |
| Female (ref. male) | 2.85 | 0.13 | <.001 | 2.85 | 0.13 | <.001 | 2.84 | 0.13 | <.001 |
| Age 14 (ref. age ≤13 years) | 1.56 | 0.16 | .007 | 1.55 | 0.16 | .008 | 1.53 | 0.16 | .009 |
| Age 15 (ref. age ≤13 years) | 1.30 | 0.17 | .134 | 1.27 | 0.17 | .162 | 1.24 | 0.17 | .208 |
| Age 16 (ref. age ≤13 years) | 1.41 | 0.18 | .056 | 1.37 | 0.18 | .074 | 1.33 | 0.18 | .111 |
| Age ≥17 (ref. age ≤13 years) | 1.06 | 0.20 | .784 | 1.02 | 0.20 | .917 | 0.98 | 0.20 | .916 |
| Minor urban (ref. main urban) | 1.16 | 0.17 | .396 | 1.13 | 0.17 | .482 | 1.09 | 0.17 | .610 |
| Rural (ref. main urban) | 1.12 | 0.18 | .528 | 1.10 | 0.18 | .585 | 1.07 | 0.18 | .680 |
| NZDep3–4 (ref. NZDep1–2) | 0.97 | 0.21 | .888 | 0.96 | 0.21 | .854 | 0.98 | 0.21 | .941 |
| NZDep5–6 (ref. NZDep1–2) | 1.25 | 0.20 | .250 | 1.23 | 0.20 | .287 | 1.30 | 0.20 | .183 |
| NZDep7–8 (ref. NZDep1–2) | 1.09 | 0.20 | .666 | 1.08 | 0.20 | .716 | 1.17 | 0.20 | .443 |
| NZDep9–10 (ref. NZDep1–2) | 1.71 | 0.19 | .004 | 1.66 | 0.19 | .007 | 1.81 | 0.19 | .002 |
| Māori (ref. European[b]) | 1.83 | 0.29 | .037 | 2.26 | 0.15 | <.001 | 1.90 | 0.17 | <.001 |
| Pacific (ref. European[b]) | 2.41 | 0.22 | <.001 | 2.58 | 0.18 | <.001 | 2.04 | 0.17 | <.001 |
| Asian (ref. European[b]) | 1.10 | 0.25 | .693 | 1.43 | 0.20 | .076 | 1.10 | 0.20 | .639 |
| Other (ref. European[b]) | 1.50 | 0.43 | .347 | 2.00 | 0.23 | .003 | 2.12 | 0.28 | .008 |
| Māori/European (ref. European[b]) | 1.83 | 0.18 | .001 | | - | | | - | |
| Pacific/European (ref. European[b]) | 2.63 | 0.23 | <.001 | | - | | | - | |
| Asian/European (ref. European[b]) | 2.33 | 0.32 | .008 | | - | | | - | |
| Māori/Pacific (ref. European[b]) | 3.29 | 0.40 | .003 | | - | | | - | |
| Māori/Pacific/European (ref. European[b]) | 3.38 | 0.34 | <.001 | | - | | | - | |
| 2 groups NEI (ref. European[b]) | 2.44 | 0.20 | <.001 | | - | | | - | |
| ≥3 groups NEI (ref. European[b]) | 3.44 | 0.23 | <.001 | | - | | | - | |
| Can't choose one ethnic group (ref. European[b]) | | - | | | - | | 1.66 | 0.43 | .238 |

*Note*. NEI = not elsewhere included.

[a]0 = no suicide attempt in the past 12 months; 1 = suicide attempt in the past 12 months.

[b]For the sole/combination and administrative-prioritisation models, the referent is *sole* European; for the self-prioritisation model, the referent is *self-prioritised* European.

# References

Achenbach, T. M., Rescorla, L. A., & Ivanova, M. Y. (2012). International epidemiology of child and adolescent psychopathology I: Diagnoses, dimensions, and conceptual issues. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(12), 1261–1272. https://doi.org/10.1016/j.jaac.2012.09.010

Adkins, D. E., Wang, V., Dupre, M. E., van den Oord, E. J. C. G., & Elder, G. H. (2009). Structure and stress: Trajectories of depressive symptoms across adolescence and young adulthood. *Social Forces*, *88*(1), 31–60. https://doi.org/10.1353/sof.0.0238

Anderson, E. R., & Mayes, L. C. (2010). Race/ethnicity and internalizing disorders in youth: A review. *Clinical Psychology Review*, *30*(3), 338–348. https://doi.org/10.1016/j.cpr.2009.12.008

Annamma, S. A., Jackson, D. D., & Morrison, D. (2017). Conceptualizing color-evasiveness: Using dis/ability critical race theory to expand a color-blind racial ideology in education and society. *Race Ethnicity and Education*, *20*(2), 147–162. https://doi.org/10.1080/13613324.2016.1248837

Aspinall, P. J. (2018a). Measuring the health patterns of the 'mixed/multiple' ethnic group in Britain: Data quality problems, reporting issues, and implications for policy. *International Journal of Social Research Methodology*, *21*(3), 359–371. https://doi.org/10.1080/13645579.2017.1399623

Aspinall, P. J. (2018b). What kind of mixed race/ethnicity data is needed for the 2020/21 global population census round: The cases of the UK, USA, and Canada. *Ethnic and Racial Studies*, *41*(11), 1990–2008. https://doi.org/10.1080/01419870.2017.1346267

Aspinall, P. J., & Song, M. (2013). *Mixed race identities*. Palgrave Macmillan.

Ataera-Minster, J., & Trowland, H. (2018). *Te Kaveinga: Mental health and wellbeing of*

*Pacific Peoples*. Health Promotion Agency.

Atatoa Carr, P., Bandara, D., Berry, S., Kingi, T., Grant, C. C., & Morton, S. (2017). Ethnic identification complexity across generations: Evidence from Growing Up in New Zealand. *New Zealand Population Review*, *43*, 35–61.

Atkinson, J., Salmond, C., & Crampton, P. (2014). *NZDep2013 Index of Deprivation*. University of Otago.

Baker, M. G., Barnard, L. T., Kvalsvig, A., Verrall, A., Zhang, J., Keall, M., Wilson, N., Wall, T., & Howden-Chapman, P. (2012). Increasing incidence of serious infectious diseases and inequalities in New Zealand: A national epidemiological study. *Lancet*, *379*(9821), 1112–1119. https://doi.org/10.1016/S0140-6736(11)61780-7

Balestra, C., & Fleischer, L. (2018). *Diversity statistics in the OECD: How do OECD countries collect data on ethnic, racial and indigenous identity?* (OECD Statistics Working Papers, No. 2018/09). OECD Publishing.

Bécares, L., Cormack, D., & Harris, R. (2013). Ethnic density and area deprivation: Neighbourhood effects on Māori health and racial discrimination in Aotearoa/New Zealand. *Social Science & Medicine*, *88*, 76–82. https://doi.org/10.1016/j.socscimed.2013.04.007

Benner, A. D., Wang, Y., Shen, Y., Boyle, A. E., Polk, R., & Cheng, Y.-P. (2018). Racial/ethnic discrimination and well-being during adolescence: A meta-analytic review. *American Psychologist*, *73*(7), 855–883. https://doi.org/10.1037/amp0000204

Bhopal, R. (2004). Glossary of terms relating to ethnicity and race: For reflection and debate. *Journal of Epidemiology & Community Health*, *58*(6), 441–445. https://doi.org/10.1136/jech.2003.013466

Blakely, T., Tobias, M., Robson, B., Ajwani, S., Bonné, M., & Woodward, A. (2005). Widening ethnic mortality disparities in New Zealand 1981–99. *Social Science &*

*Medicine*, *61*(10), 2233–2251. https://doi.org/10.1016/j.socscimed.2005.02.011

Boccagni, P. (2015). (Super)diversity and the migration–social work nexus: A new lens on the field of access and inclusion? *Ethnic and Racial Studies*, *38*(4), 608–620. https://doi.org/10.1080/01419870.2015.980291

Bokor-Billmann, T., Langan, E. A., & Billmann, F. (2020). The reporting of race and/or ethnicity in the medical literature: A retrospective bibliometric analysis confirmed room for improvement. *Journal of Clinical Epidemiology*, *119*, 1–6. https://doi.org/10.1016/j.jclinepi.2019.11.005

Bonilla-Silva, E., & Zuberi, T. (2008). Toward a definition of White logic and White methods. In T. Zuberi & E. Bonilla-Silva (Eds.), *White logic, White methods: Racism and methodology* (pp. 3–27). Rowman & Littlefield.

Borrell, B. (2005). Living in the city ain't so bad: Cultural identity for young Maori in South Auckland. In J. H. Liu, T. McCreanor, T. McIntosh, & T. Teaiwa (Eds.), *New Zealand identities: Departures and destinations* (pp. 369–400). Victoria University Press.

Boven, N., Exeter, D., Sporle, A., & Shackleton, N. (2020). The implications of different ethnicity categorisation methods for understanding outcomes and developing policy in New Zealand. *Kōtuitui: New Zealand Journal of Social Sciences Online*, *15*(1), 123–139. https://doi.org/10.1080/1177083X.2019.1657912

Bowleg, L. (2012). The problem with the phrase women and minorities: Intersectionality—an important theoretical framework for public health. *American Journal of Public Health*, *102*(7), 1267–1273. https://doi.org/10.2105/AJPH.2012.300750

Bratter, J. L. (2007). Will "multiracial" survive to the next generation?: The racial classification of children of multiracial parents. *Social Forces*, *86*(2), 821–849. https://doi.org/10.1093/sf/86.2.821

Bratter, J. L. (2018). Multiracial identification and racial gaps: A work in progress. *Annals of*

*the American Academy of Political and Social Science*, *677*, 69–80. https://doi.org/10.1177/0002716218758622

Bratter, J. L., & Gorman, B. K. (2011). Does multiracial matter? A study of racial disparities in self-rated health. *Demography*, *48*(1), 127–152. https://doi.org/10.1007/s13524-010-0005-0

Broman, P. (2018). Measuring the majority: Counting Europeans in the New Zealand census. *New Zealand Sociology*, *33*(3), 83–107.

Brunsma, D. L. (2005). Interracial families and the racial identification of mixed-race children: Evidence from the Early Childhood Longitudinal Study. *Social Forces*, *84*(2), 1131–1157. https://doi.org/10.1353/sof.2006.0007

Callister, P., & Blakely, T. (2004). *Ethnic classification, intermarriage, and mortality: Some methodological issues in relation to ethnic comparisons in Aotearoa/New Zealand*. Wellington School of Medicine and Health Sciences.

Callister, P., Didham, R., Potter, D., & Blakely, T. (2007). Measuring ethnicity in New Zealand: Developing tools for health outcomes analysis. *Ethnicity and Health*, *12*(4), 299–320. https://doi.org/10.1080/13557850701300699

Campbell, M. E. (2007). Thinking outside the (black) box: Measuring black and multiracial identification on surveys. *Social Science Research*, *36*(3), 921–944. https://doi.org/10.1016/j.ssresearch.2006.07.001

Campbell, M. E., & Eggerling-Boeck, J. (2006). "What about the children?" The psychological and social well-being of multiracial adolescents. *The Sociological Quarterly*, *47*(1), 147–173. https://doi.org/10.1111/j.1533-8525.2006.00041.x

Carter, K. N., Hayward, M., Blakely, T., & Shaw, C. (2009). How much and for whom does self-identified ethnicity change over time in New Zealand? Results from a longitudinal study. *Social Policy Journal of New Zealand*, *36*, 32–45.

Charmaraman, L., Woo, M., Quach, A., & Erkut, S. (2014). How have researchers studied multiracial populations? A content and methodological review of 20 years of research. *Cultural Diversity and Ethnic Minority Psychology*, *20*(3), 336–352. https://doi.org/10.1037/a0035437

Chatterjee, S., & Simonoff, J. S. (2013). *Handbook of regression analysis*. John Wiley & Sons.

Cheng, S., & Lively, K. J. (2009). Multiracial self-identification and adolescent outcomes: A social psychological approach to the marginal man theory. *Social Forces*, *88*(1), 61–98. https://doi.org/10.1353/sof.0.0243

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*(22), 3127–3131. https://doi.org/10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M

Clark, T. C., Fleming, T., Bullen, P., Crengle, S., Denny, S., Dyson, B., Fortune, S., Peiris-John, R., Robinson, E., Rossen, F., Sheridan, J., Teevale, T., Utter, J., & The Adolescent Health Research Group. (2013). *Youth'12 prevalence tables: The health and wellbeing of New Zealand secondary school students in 2012*. The University of Auckland.

Clark, T. C., Le Grice, J., Moselen, E., Fleming, T., Crengle, S., Tiatia-Seath, J., & Lewycka, S. (2018). Health and wellbeing of Māori secondary school students in New Zealand: Trends between 2001, 2007 and 2012. *Australian and New Zealand Journal of Public Health*, *42*(6), 553–561. https://doi.org/10.1111/1753-6405.12839

Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic review and meta-analysis: Adolescent depression and long-term psychosocial outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, *58*(1), 72–79. https://doi.org/10.1016/j.jaac.2018.07.896

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence

Erlbaum Associates.

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

Cokley, K. (2007). Critical issues in the measurement of ethnic and racial identity: A referendum on the state of the field. *Journal of Counseling Psychology*, *54*(3), 224–234. https://doi.org/10.1037/0022-0167.54.3.224

Cokley, K., & Awad, G. H. (2013). In defense of quantitative methods: Using the "master's tools" to promote social justice. *Journal for Social Action in Counseling & Psychology*, *5*(2), 26–41. https://doi.org/10.33043/JSACP.5.2.26-41

Comstock, R. D. (2004). Four-year review of the use of race and ethnicity in epidemiologic and public health research. *American Journal of Epidemiology*, *159*(6), 611–619. https://doi.org/10.1093/aje/kwh084

Cormack, D., Reid, P., & Kukutai, T. (2019). Indigenous data and health: Critical approaches to 'race'/ethnicity and Indigenous data governance. *Public Health*, *172*, 116–118. https://doi.org/10.1016/j.puhe.2019.03.026

Cormack, D., & Robson, C. (2010). *Classification and output of multiple ethnicities: Considerations for Māori health*. Te Rōpū Rangahau Hauora a Eru Pōmare.

Craven, M. R., Kia, L., O'Dwyer, L. C., Stern, E., Taft, T. H., & Keefer, L. (2018). Systematic review: Methodological flaws in racial/ethnic reporting for gastroesophageal reflux disease. *Diseases of the Esophagus*, *31*(3), 1–12. https://doi.org/10.1093/dote/dox154

Crengle, S., Robinson, E., Ameratunga, S., Clark, T., & Raphael, D. (2012). Ethnic discrimination prevalence and associations with health outcomes: Data from a nationally representative cross-sectional survey of secondary school students in New Zealand.

*BMC Public Health*, *12*(1), 45. https://doi.org/10.1186/1471-2458-12-45

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, *1989*(139), 139–167. https://doi.org/10.4324/9780429499142-5

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). SAGE Publications.

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, *60*(2), 170–180. https://doi.org/10.1037/0003-066X.60.2.170

Curtis, E., Wright, C., & Wall, M. (2005). The epidemiology of breast cancer in Maori women in Aotearoa New Zealand: Implications for ethnicity data analysis. *New Zealand Medical Journal*, *118*(1209), 1–13.

Davis, P., Lay-Yee, R., Dyall, L., Briant, R., Sporle, A., Brunt, D., & Scott, A. (2006). Quality of hospital care for Māori patients in New Zealand: Retrospective cross-sectional assessment. *Lancet*, *367*(9526), 1920–1925. https://doi.org/10.1016/S0140-6736(06)68847-8

de Jong, V. M. T., Eijkemans, M. J. C., van Calster, B., Timmerman, D., Moons, K. G. M., Steyerberg, E. W., & van Smeden, M. (2019). Sample size considerations and predictive performance of multinomial logistic prediction models. *Statistics in Medicine*, *38*(9), 1601–1619. https://doi.org/10.1002/sim.8063

Denton, N. A., & Deane, G. D. (2010). Researching race and ethnicity: Methodological issues. In P. H. Collins & J. Solomos (Eds.), *The SAGE handbook of race and ethnic studies* (pp. 67–89). SAGE Publications.

Department of Statistics. (1993). *New Zealand standard classification of ethnicity*.

Didham, R. (2005). *Understanding and working with ethnicity data*. Statistics New Zealand.

Didham, R. (2016). Ethnic mobility in the New Zealand census, 1981– 2013: A preliminary look. *New Zealand Population Review*, *42*, 27–42.

Didham, R., & Callister, P. (2012). The effect of ethnic prioritisation on ethnic health analysis: A research note. *New Zealand Medical Journal*, *125*(1359), 58–66.

Dwyer, S. C., & Buckle, J. L. (2009). The space between: On being an insider-outsider in qualitative research. *International Journal of Qualitative Methods*, *8*(1), 54–63. https://doi.org/10.1177/160940690900800105

Education Counts. (2014). *Ethnic group codes*. https://www.educationcounts.govt.nz/data-services/collecting-information/code-sets-and-classifications/ethnic_group_codes

Espiner, E., Paine, S.-J., Weston, M., & Curtis, E. (2021). Barriers and facilitators for Māori in accessing hospital services in Aotearoa New Zealand. *New Zealand Medical Journal*, *134*(1546), 47–58.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE Publications.

Finlay, L. (2002). "Outing" the researcher: The provenance, process, and practice of reflexivity. *Qualitative Health Research*, *12*(4), 531–545. https://doi.org/10.1177/104973202129120052

Fisher, S., Reynolds, J. L., Hsu, W.-W., Barnes, J., & Tyler, K. (2014). Examining multiracial youth in context: Ethnic identity development and mental health outcomes. *Journal of Youth and Adolescence*, *43*(10), 1688–1699. https://doi.org/10.1007/s10964-014-0163-2

Fleming, T., Peiris-John, R., Crengle, S., Archer, D., Sutcliffe, K., Lewycka, S., & T., C. (2020). *Youth19 Rangatahi Smart Survey, initial findings: Introduction and methods*. The Youth19 Research Group, The University of Auckland, and Victoria University of Wellington.

Fleming, T., Tiatia-Seath, J., Peiris-John, R., Sutcliffe, K., Archer, D., Bavin, L., Crengle, S.,

& Clark, T. (2020). *Youth19 Rangatahi Smart Survey, initial findings: Hauora hinengaro/emotional and mental health*. The Youth19 Research Group, The University of Auckland, and Victoria University of Wellington.

Fortune, S., Watson, P., Robinson, E., Fleming, T., Merry, S., & Denny, S. (2010). *Youth'07: The health and wellbeing of secondary school students in New Zealand: Suicide behaviours and mental health in 2001 and 2007*. The University of Auckland.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). SAGE Publications.

Gaither, S. E. (2015). "Mixed" results: Multiracial research and identity explorations. *Current Directions in Psychological Science*, *24*(2), 114–119. https://doi.org/10.1177/0963721414558115

Garcia, N. M., López, N., & Vélez, V. N. (2018). QuantCrit: Rectifying quantitative methods through critical race theory. *Race Ethnicity and Education*, *21*(2), 149–157. https://doi.org/10.1080/13613324.2017.1377675

Gillborn, D., Warmington, P., & Demack, S. (2018). QuantCrit: Education, policy, 'big data' and principles for a critical race theory of statistics. *Race Ethnicity and Education*, *21*(2), 158–179. https://doi.org/10.1080/13613324.2017.1377417

Goodman, A., & Goodman, R. (2009). Strengths and Difficulties Questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(4), 400–403. https://doi.org/10.1097/CHI.0b013e3181985068

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry*, *7*(3), 125–130. https://doi.org/10.1007/s007870050057

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual*

*Review of Psychology*, *60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Gravlee, C. C. (2009). How race becomes biology: Embodiment of social inequality. *American Journal of Physical Anthropology*, *139*(1), 47–57. https://doi.org/10.1002/ajpa.20983

Grieco, E. M. (2002). An evaluation of bridging methods using race data from Census 2000. *Population Research and Policy Review*, *21*, 91–107.

Gullickson, A., & Morning, A. (2011). Choosing race: Multiracial ancestry and identification. *Social Science Research*, *40*(2), 498–512. https://doi.org/10.1016/j.ssresearch.2010.12.010

Hancock, A.-M. (2007). When multiplication doesn't equal quick addition: Examining intersectionality as a research paradigm. *Perspectives on Politics*, *5*(1), 63–79. https://doi.org/10.1017/S1537592707070065

Harris, D. R., & Sim, J. J. (2002). Who is multiracial? Assessing the complexity of lived race. *American Sociological Review*, *67*(4), 614–627. https://doi.org/10.2307/3088948

Harris, R., Cormack, D., Tobias, M., Yeh, L.-C., Talamaivao, N., Minster, J., & Timutimu, R. (2012). The pervasive effects of racism: Experiences of racial discrimination in New Zealand over time and associations with multiple health domains. *Social Science & Medicine*, *74*(3), 408–415. https://doi.org/10.1016/j.socscimed.2011.11.004

Heck, K. E., Parker, J. D., McKendry, C. J., & Chávez, G. F. (2003). Mind the gap: Bridge methods to allocate multiple-race mothers in trend analyses of birth certificate data. *Maternal and Child Health Journal*, *7*(1), 65–70. https://doi.org/10.1023/a:1022597702856

Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection – A review and recommendations for the practicing statistician. *Biometrical Journal*, *60*(3), 431–449.

https://doi.org/10.1002/bimj.201700067

Herman, M. (2004). Forced to choose: Some determinants of racial identification in multiracial adolescents. *Child Development*, *75*(3), 730–748. https://doi.org/10.1111/j.1467-8624.2004.00703.x

Herman, M. R. (2011). Methodology and measurement in the study of multiracial Americans: Identity, classification, and perceptions. *Sociology Compass*, *5*(7), 607–617. https://doi.org/10.1111/j.1751-9020.2011.00388.x

Hirschman, C. (2004). The origins and demise of the concept of race. *Population and Development Review*, *30*(3), 385–415. https://doi.org/10.1111/j.1728-4457.2004.00021.x

Hobbs, M. R., Atatoa Carr, P., Fa'Alili-Fidow, J., Pillai, A., Morton, S. M. B., & Grant, C. C. (2019). How differing methods of ascribing ethnicity and socio-economic status affect risk estimates for hospitalisation with infectious disease. *Epidemiology and Infection*, *147*. https://doi.org/10.1017/S0950268818002935

Hong Kong University Public Opinion Programme. (2019). *HKU POP final farewell: Rift widens between Chinese and Hongkong identities, national pride plunges to one in four*. https://www.hkupop.hku.hk/english/release/release1594.html

Hudson, M., Milne, M., Reynolds, P., Russell, K., & Smith, B. (2010). *Te Ara Tika: Guidelines for Māori research ethics: A framework for researchers and ethics committee members*. Health Research Council of New Zealand.

Ingram, D. D., Parker, J. D., Schenker, N., Weed, J. A., Hamilton, B., Arias, E., & Madans, J. H. (2003). United States Census 2000 population with bridged race categories. *Vital and Health Statistics*, *2*(135).

Jackson, M. (2003). The part-Maori syndrome. *Mana*, *52*, 62.

Jones, C. P. (2000). Levels of racism: A theoretic framework and a gardener's tale. *American Journal of Public Health*, *90*(8), 1212–1215.

Jones, N. A., & Bullock, J. (2012). *The two or more races population: 2010* (2010 Census Briefs). U.S. Census Bureau.

Jones, N., Marks, R., Ramirez, R., & Ríos-Vargas, M. (2021). *Improved race and ethnicity measures reveal U.S. population is much more multiracial*. U.S. Census Bureau. https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html

Jorde, L. B., & Wooding, S. P. (2004). Genetic variation, classification and "race." *Nature Genetics*, *36*, S28–S33. https://doi.org/10.1034/ng1435

Kertzer, D. I., & Arel, D. (2002). *Census and identity: The politics of race, ethnicity, and language in national censuses*. Cambridge University Press.

Kessler, R. C., Avenevoli, S., Costello, E. J., Georgiades, K., Green, J. G., Gruber, M. J., He, J., Koretz, D., McLaughlin, K. A., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Merikangas, K. R. (2012). Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, *69*(4), 372–380. https://doi.org/10.1001/archgenpsychiatry.2011.160

Khanna, N. (2012). Multiracial Americans: Racial identity choices and implications for the collection of race data. *Sociology Compass*, *6*(4), 316–331. https://doi.org/10.1111/j.1751-9020.2011.00454.x

Khawaja, M., Boddington, B., & Didham, R. (2000). *Growing ethnic diversity in New Zealand and its implications for measuring differentials in fertility and mortality*. Statistics New Zealand.

Kincheloe, J. L., McLaren, P., Steinberg, S. R., & Monzó, L. D. (2018). Critical pedagogy and qualitative research: Advancing the bricolage. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (5th ed., pp. 235–260). SAGE

Publications.

Kingi, T. K. R. (2007). The Treaty of Waitangi: A framework for Māori health development. *New Zealand Journal of Occupational Therapy*, *54*(1), 4–10. https://doi.org/10.3316/informit.304488051134401

Krieger, N. (2014). Discrimination and health inequities. *International Journal of Health Services*, *44*(4), 643–710. https://doi.org/10.2190/HS.44.4.b

Kukutai, T. (2004). The problem of defining an ethnic group for public policy: Who is Maori and why does it matter. *Social Policy Journal of New Zealand*, *23*(23), 86–108.

Kukutai, T., & Callister, P. (2009). A "main" ethnic group? Ethnic self-prioritisation among New Zealand youth. *Social Policy Journal of New Zealand*, *36*, 16–31. https://doi.org/10.1787/89bae654-en

Kukutai, T., & Didham, R. (2009). In search of ethnic New Zealanders: National naming in the 2006 census. *Social Policy Journal of New Zealand*, *36*, 46–62.

Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. Australian National University Press.

Kukutai, T., & Walter, M. (2015). Recognition and indigenizing official statistics: Reflections from Aotearoa New Zealand and Australia. *Statistical Journal of the IAOS*, *31*(2), 317–326. https://doi.org/10.3233/sji-150896

Lachowsky, N. J., Saxton, P. J. W., Dickson, N. P., Hughes, A. J., Jones, R. G., Clark, T. C., Ho, E., Summerlee, A. J. S., & Dewey, C. E. (2020). Ethnicity classification systems for public health surveys: Experiences from HIV behavioural surveillance among men who have sex with men. *BMC Public Health*, *20*, 1–13. https://doi.org/10.1186/s12889-020-09517-4

Lee, C. (2009). "Race" and "ethnicity" in biomedical research: How do scientists construct and explain differences in health? *Social Science & Medicine*, *68*(6), 1183–1190.

https://doi.org/10.1016/j.socscimed.2008.12.036

Lee, S. J., Wong, N.-W. A., & Alvarez, A. N. (2009). The model minority and the perpetual foreigner: Stereotypes of Asian Americans. In N. Tewari & A. N. Alvarez (Eds.), *Asian American psychology: Current perspectives* (pp. 69–84). Psychology Press.

Lee, S. M. (2001). *Using the new racial categories in the 2000 Census* (KIDS COUNT/PRB Reports on Census 2000). Annie E. Casey Foundation and Population Reference Bureau.

Liebler, C. A., & Halpern-Manners, A. (2008). A practical approach to using multiple-race response data: A bridging method for public-use microdata. *Demography*, *45*(1), 143–155. https://doi.org/10.1353/dem.2008.0004

Liebler, C. A., Porter, S. R., Fernandez, L. E., Noon, J. M., & Ennis, S. R. (2017). America's churning races: Race and ethnicity response changes between Census 2000 and the 2010 Census. *Demography*, *54*(1), 259–284. https://doi.org/10.1007/s13524-016-0544-0

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2018). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (5th ed., pp. 108–150). SAGE Publications.

Luk, G. C. (2017). Straddling the handover: Colonialism and decolonization in British and PRC Hong Kong. In G. C. Luk (Ed.), *From a British to a Chinese colony? Hong Kong before and after the 1997 Handover* (pp. 1–50). Institute of East Asian Studies.

Maduka, R. C., Broderick, M., White, E. M., Ballouz, D., Sandhu, H., Kwakye, G., Chen, H., & Sandhu, G. (2021). The reporting of race and ethnicity in surgery literature. *JAMA Surgery*, *156*(11), 1036–1041. https://doi.org/10.1001/jamasurg.2021.37520.1001/jamasurg.2021.3752

Mayhew, M. J., & Simonoff, J. S. (2015). Non-white, no more: Effect coding as an alternative to dummy coding with implications for higher education researchers. *Journal*

*of College Student Development*, *56*(2), 170–175. https://doi.org/10.1353/csd.2015.0019

Mays, V. M., Ponce, N. A., Washington, D. L., & Cochran, S. D. (2003). Classification of race and ethnicity: Implications for public health. *Annual Review of Public Health*, *24*(1), 83–110. https://doi.org/10.1146/annurev.publhealth.24.100901.140927

Meissner, F. (2015). Migration in migration-related diversity? The nexus between superdiversity and migration studies. *Ethnic and Racial Studies*, *38*(4), 556–567. https://doi.org/10.1080/01419870.2015.970209

Mertens, D. M. (2019). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (5th ed.). SAGE Publications.

Ministry of Health. (2008). *Presenting ethnicity: Comparing prioritised and total response ethnicity in descriptive analyses of New Zealand Health Monitor surveys*.

Ministry of Health. (2017). *Ethnicity data protocols* (HISO 10001:2017).

Ministry of Social Development. (2016). *The Social Report 2016*.

Mojtabai, R., Olfson, M., & Han, B. (2016). National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics*, *138*(6), e20161878–e20161878. https://doi.org/10.1542/peds.2016-1878

Mok, C., & Defranco, A. L. (2000). Chinese cultural values: Their implications for travel and tourism marketing. *Journal of Travel & Tourism Marketing*, *8*(2), 99–114. https://doi.org/10.1300/J073v08n02_07

Møllersen, S., & Holte, A. (2008). Ethnicity as a variable in mental health research: A systematic review of articles published 1990–2004. *Nordic Journal of Psychiatry*, *62*(4), 322–328. https://doi.org/10.1080/08039480802044414

Moore, D. B. (2020). Reporting of race and ethnicity in the ophthalmology literature in 2019. *JAMA Ophthalmology*, *138*(8), 903–906.

https://doi.org/10.1001/jamaophthalmol.2020.2107

Morning, A. (2008). Ethnic classification in global perspective: A cross-national survey of the 2000 census round. *Population Research and Policy Review*, *27*(2), 239–272. https://doi.org/10.1007/s11113-007-9062-5

Morton, S. M. B., Atatoa Carr, P. E., Grant, C. C., Robinson, E. M., Bandara, D. K., Bird, A., Ivory, V. C., Kingi, T. K. R., Liang, R., Marks, E. J., Perese, L. M., Peterson, E. R., Pryor, J. E., Reese, E., Schmidt, J. M., Waldie, K. E., & Wall, C. (2013). Cohort profile: Growing Up in New Zealand. *International Journal of Epidemiology*, *42*(1), 65–75. https://doi.org/10.1093/ije/dyr206

Morton, S. M. B., Ramke, J., Kinloch, J., Grant, C. C., Atatoa Carr, P., Leeson, H., Lee, A. C. L., & Robinson, E. (2015). Growing Up in New Zealand cohort alignment with all New Zealand births. *Australian and New Zealand Journal of Public Health*, *39*(1), 82–87. https://doi.org/10.1111/1753-6405.12220

Moubarac, J.-C. (2013). Persisting problems related to race and ethnicity in public health and epidemiology research. *Revista de Saúde Pública*, *47*(1), 105–116. https://doi.org/10.1590/S0034-89102013000100014

Moya, P. M. L., & Markus, H. R. (2010). Doing race: An introduction. In P. M. L. Moya & H. R. Markus (Eds.), *Doing race: 21 essays for the 21st century* (pp. 1–102). W. W. Norton & Company.

Nandi, A., & Platt, L. (2012). How diverse is the UK? In S. McFall, N. Buck, H. Laurie, C. Garrington, & V. Nolan (Eds.), *Understanding Society: Findings 2012* (pp. 13–14). Institute for Social and Economic Research, University of Essex.

Nishina, A., Bellmore, A., Witkow, M. R., & Nylund-Gibson, K. (2010). Longitudinal consistency of adolescent ethnic identification across varying school ethnic contexts. *Developmental Psychology*, *46*(6), 1389–1401. https://doi.org/10.1037/a0020728

Nishina, A., & Witkow, M. R. (2020). Why developmental researchers should care about biracial, multiracial, and multiethnic youth. *Child Development Perspectives*, *14*(1), 21–27. https://doi.org/10.1111/cdep.12350

Office of Management and Budget. (2000). *Provisional guidance on the implementation of the 1997 standards for the collection of federal data on race and ethnicity*. Executive Office of the President.

Omi, M., & Winant, H. (2015). *Racial formation in the United States* (3rd ed.). Routledge.

Orange, C. (2011). *The Treaty of Waitangi*. Bridget Williams Books.

Parker, J. D., & Makuc, D. M. (2002). Methodologic implications of allocating multiple-race data to single-race categories. *Health Services Research*, *37*(1), 201–213. https://doi.org/10.1111/1475-6773.00091

Peiris-John, R., Bavin, L., Kang, K., Dizon, L., Lewycka, S., Ameratunga, S., Clark, T., & Fleming, T. (2022). Factors predicting forgone healthcare among Asian adolescents in New Zealand: Unmasking variations in aggregate data. *New Zealand Medical Journal*, *135*(1549), 63–80.

Perez, A. D. (2006). *Choosing a race: The effects of proxy reporting on the identification of mixed race children*. Population Studies Center, University of Michigan.

Perez, A. D., & Hirschman, C. (2009). The changing racial and ethnic composition of the US population: Emerging American identities. *Population and Development Review*, *35*(1), 1–51. https://doi.org/10.1111/j.1728-4457.2009.00260.x

Phinney, J. S., & Ong, A. D. (2007). Conceptualization and measurement of ethnic identity: Current status and future directions. *Journal of Counseling Psychology*, *54*(3), 271–281. https://doi.org/10.1037/0022-0167.54.3.271

Punch, K. F. (2014). *Introduction to social research: Quantitative and qualitative approaches* (3rd ed.). SAGE Publications.

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org

Rallu, J.-L., Piché, V., & Simon, P. (2006). Demography and ethnicity: An ambiguous relationship. In G. Caselli, J. Vallin, & G. Wunsch (Eds.), *Demography: Analysis and synthesis. A treatise in population studies* (Vol. 3, pp. 531–549). Academic Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). SAGE Publications.

Research Data Alliance International Indigenous Data Sovereignty Interest Group. (2019). *CARE principles for Indigenous data governance*. The Global Indigenous Data Alliance.

Rios-Aguilar, C. (2014). The changing context of critical quantitative inquiry. *New Directions for Institutional Research*, *2013*(158), 95–107. https://doi.org/10.1002/ir.20048

Ro, H. K., & Bergom, I. (2020). Expanding our methodological toolkit: Effect coding in critical quantitative studies. *New Directions for Student Services*, *2020*(169), 87–97. https://doi.org/10.1002/ss.20347

Rocha, Z. L. (2012). (Mixed) racial formation in Aotearoa/New Zealand: Framing biculturalism and 'mixed race' through categorisation. *Kōtuitui: New Zealand Journal of Social Sciences Online*, *7*(1), 1–13. https://doi.org/10.1080/1177083X.2012.670650

Rocha, Z. L., & Aspinall, P. J. (2020). Introduction: Measuring mixedness around the world. In Z. L. Rocha & P. J. Aspinall (Eds.), *The Palgrave international handbook of mixed racial and ethnic classification* (pp. 1–26). Palgrave Macmillan.

Rocha, Z. L., & Webber, M. (2018). *Mana tangatarua: Mixed heritages, ethnic identity and biculturalism in Aotearoa/New Zealand*. Routledge.

Roth, W. D. (2005). The end of the one-drop rule? Labeling of multiracial children in black

intermarriages. *Sociological Forum*, *20*(1), 35–67. https://doi.org/10.1007/s11206-005-1897-0

Roth, W. D. (2016). The multiple dimensions of race. *Ethnic and Racial Studies*, *39*(8), 1310–1338. https://doi.org/10.1080/01419870.2016.1140793

Rutkowski, R. E., Salemi, J. L., Tanner, J. P., Matas, J. L., & Kirby, R. S. (2017). Assessing the impact of different race-bridging algorithms on the reported rate of birth defects. *Journal of Registry Management*, *44*(4), 145–156.

Sablan, J. R. (2019). Can you really measure that? Combining critical race theory and quantitative methods. *American Educational Research Journal*, *56*(1), 178–203. https://doi.org/10.3102/0002831218798325

Salmond, C., & Crampton, P. (2012). Measuring socioeconomic position in New Zealand. *Journal of Primary Health Care*, *4*(4), 271–280. https://doi.org/10.1071/HC12280

Sanchez, D. T. (2010). How do forced-choice dilemmas affect multiracial people? The role of identity autonomy and public regard in depressive symptoms. *Journal of Applied Social Psychology*, *40*(7), 1657–1677. https://doi.org/10.1111/j.1559-1816.2010.00634.x

Sanchez, D. T., Gaither, S. E., Albuja, A. F., & Eddy, Z. (2020). How policies can address multiracial stigma. *Policy Insights from the Behavioral and Brain Sciences*, *7*(2), 115–122. https://doi.org/10.1177/2372732220943906

Sankar, P., Cho, M. K., Monahan, K., & Nowak, K. (2015). Reporting race and ethnicity in genetics research: Do journal recommendations or resources matter? *Science and Engineering Ethics*, *21*(5), 1353–1366. https://doi.org/10.1007/s11948-014-9596-y

Schenker, N., & Parker, J. D. (2003). From single-race reporting to multiple-race reporting: Using imputation methods to bridge the transition. *Statistics in Medicine*, *22*(9), 1571–1587. https://doi.org/10.1002/sim.1512

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data

management in counseling psychology. *Journal of Counseling Psychology*, *57*(1), 1–10. https://doi.org/10.1037/a0018082

Shih, M., & Sanchez, D. T. (2005). Perspectives and research on the positive and negative implications of having multiple racial identities. *Psychological Bulletin*, *131*(4), 569–591. https://doi.org/10.1037/0033-2909.131.4.569

Simon, P. (2012). Collecting ethnic statistics in Europe: A review. *Ethnic and Racial Studies*, *35*(8), 1366–1391. https://doi.org/10.1080/01419870.2011.607507

Skinner, A. L., Perry, S. P., & Gaither, S. (2020). Not quite monoracial: Biracial stereotypes explored. *Personality and Social Psychology Bulletin*, *46*(3), 377–392. https://doi.org/10.1177/0146167219858344

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist*, *60*(1), 16–26. https://doi.org/10.1037/0003-066X.60.1.16

Smith, L. T. (2021). *Decolonizing methodologies: Research and Indigenous Peoples* (3rd ed.). Zed Books.

Stage, F. K. (2007). Answering critical questions using quantitative data. *New Directions for Institutional Research*, *2007*(133), 5–16. https://doi.org/10.1002/ir.200

Stage, F. K., & Wells, R. S. (2014). Critical quantitative inquiry in context. *New Directions for Institutional Research*, *2013*(158), 1–7. https://doi.org/10.1002/ir.20041

Statistics New Zealand. (2004). *Report of the review of the measurement of ethnicity*.

Statistics New Zealand. (2005). *Statistical standard for ethnicity*.

Statistics New Zealand. (2013). *2013 Census QuickStats about Māori*.

Statistics New Zealand. (2014a). *2013 Census meshblock dataset*.

Statistics New Zealand. (2014b). *2013 Census Quickstats about culture and identity*.

Statistics New Zealand. (2017). *Statistical standard for geographic areas 2018*.

Statistics New Zealand. (2020a). *2018 Census ethnic group summaries*. https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries

Statistics New Zealand. (2020b). *Ethnicity standard classification: Findings from public consultation November 2019*. https://www.stats.govt.nz/consultations/ethnicity-standard-classification-findings-from-public-consultation-november-2019

Subramanian, S. (2009). Methods and approaches in using secondary data sources to study race and ethnicity factors. In M. Verma (Ed.), *Cancer epidemiology* (Vol. 1, pp. 227–237). Humana Press.

Tabb, K. M., Gavin, A. R., Smith, D. C., & Huang, H. (2019). Self-rated health among multiracial young adults in the United States: Findings from the Add Health study. *Ethnicity & Health*, *24*(5), 495–511. https://doi.org/10.1080/13557858.2017.1346175

Taua'i, E., Richards, R., & Kokaua, J. (2018). Is Pacific language ability protective of prevalence of mental disorders among Pacific peoples in New Zealand? *Pacific Health Dialog*, *21*(1), 10–16. https://doi.org/10.26635/phd.2018.902

Te Rōpū Rangahau Hauora a Eru Pōmare. (2000). Counting for nothing: Understanding the issues in monitoring disparities in health. *Social Policy Journal of New Zealand*, *14*, 1–16.

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *Lancet*, *379*(9820), 1056–1067. https://doi.org/10.1016/S0140-6736(11)60871-4

Tishkoff, S. A., & Kidd, K. K. (2004). Implications of biogeography of human populations for "race" and medicine. *Nature Genetics*, *36*(S11), S21–S27. https://doi.org/10.1038/ng1438

Townsend, S. S. M., Markus, H. R., & Bergsieker, H. B. (2009). My choice, your categories: The denial of multiracial identities. *Journal of Social Issues*, *65*(1), 185–204. https://doi.org/10.1111/j.1540-4560.2008.01594.x

Udry, J. R., Li, R. M., & Hendrickson-Smith, J. (2003). Health and behavior risks of adolescents with mixed-race identity. *American Journal of Public Health*, *93*(11), 1865–1870. https://doi.org/10.2105/AJPH.93.11.1865

United Nations. (2008). *Principles and recommendations for population and housing censuses* (Series M No. 67/Rev.2).

Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies*, *30*(6), 1024–1054. https://doi.org/10.1080/01419870701599465

Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, *165*(6), 710–718. https://doi.org/10.1093/aje/kwk052

Waitangi Tribunal. (2019). *Hauora: Report on stage one of the health services and outcomes kaupapa inquiry*.

Walter, M., & Andersen, C. (2013). *Indigenous statistics: A quantitative research methodology*. Left Coast Press.

Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond "p < 0.05." *The American Statistician*, *73*(sup1), 1–19. https://doi.org/10.1080/00031305.2019.1583913

White, K., Lawrence, J. A., Tchangalova, N., Huang, S. J., & Cummings, J. L. (2020). Socially-assigned race and health: a scoping review with global implications for population health equity. *International Journal for Equity in Health*, *19*(1), 25. https://doi.org/10.1186/s12939-020-1137-5

Williams, A. D., Clark, T. C., & Lewycka, S. (2018). The associations between cultural identity and mental health outcomes for indigenous Māori youth in New Zealand. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00319

Williams, D. R. (2018). Stress and the mental health of populations of color: Advancing our

understanding of race-related stressors. *Journal of Health and Social Behavior*, *59*(4), 466–485. https://doi.org/10.1177/0022146518814251

Woo, M., Austin, S. B., Williams, D. R., & Bennett, G. G. (2011). Reconceptualizing the measurement of multiracial status for health research in the United States. *Du Bois Review*, *8*(1), 25–36. https://doi.org/10.1017/S1742058X11000038

Wyeth, E. H., Derrett, S., Hokowhitu, B., Hall, C., & Langley, J. (2010). Rangatiratanga and Ōritetanga: Responses to the Treaty of Waitangi in a New Zealand study. *Ethnicity & Health*, *15*(3), 303–316. https://doi.org/10.1080/13557851003721194

Xu, J., Saperstein, A., Morning, A., & Iverson, S. (2021). Gender, generation, and multiracial identification in the United States. *Demography*, *58*(5), 1603–1630. https://doi.org/10.1215/00703370-9334366

Yao, E. S., Bullen, P., Meissel, K., Tiatia, J., Fleming, T., & Clark, T. C. (2022). Effects of ethnic classification on substantive findings in adolescent mental health outcomes. *Journal of Youth and Adolescence*, *51*, 1581–1596. https://doi.org/10.1007/s10964-022-01612-6

Yao, E. S., Meissel, K., Bullen, P., Atatoa Carr, P., Clark, T. C., & Morton, S. M. B. (2021). Classifying multiple ethnic identifications: Methodological effects on child, adolescent, and adult ethnic distributions. *Demographic Research*, *44*, 481–512. https://doi.org/10.4054/DemRes.2021.44.21

Yao, E. S., Meissel, K., Bullen, P., Clark, T. C., Atatoa Carr, P., Tiatia-Seath, J., Peiris-John, R., & Morton, S. M. B. (2022). Demographic discrepancies between administrative-prioritisation and self-prioritisation of multiple ethnic identifications. *Social Science Research*, *103*, 1–16. https://doi.org/10.1016/j.ssresearch.2021.102648

Zuberi, T. (2001). *Thicker than blood: How racial statistics lie*. University of Minnesota Press.