

DEMOGRAPHIC RESEARCH

A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 44, ARTICLE 21, PAGES 481–512

PUBLISHED 16 MARCH 2021

<https://www.demographic-research.org/Volumes/Vol44/21/>

DOI: 10.4054/DemRes.2021.44.21

Research Article

Classifying multiple ethnic identifications: Methodological effects on child, adolescent, and adult ethnic distributions

Esther S. Yao

Polly Atatoa Carr

Kane Meissel

Terryann C. Clark

Pat Bullen

Susan M.B. Morton

© 2021 Esther S. Yao et al.

This open-access work is published under the terms of the Creative Commons Attribution 3.0 Germany (CC BY 3.0 DE), which permits use, reproduction, and distribution in any medium, provided the original author(s) and source are given credit.

See <https://creativecommons.org/licenses/by/3.0/de/legalcode>.

Contents

1	Introduction	482
1.1	Multiple ethnic identifications	483
1.2	Ethnic classification methods	484
1.2.1	Retention of multiple ethnicities	486
1.2.2	Reduction of multiple ethnicities	487
1.3	The current study	488
1.3.1	Aotearoa/New Zealand context	489
1.3.2	Research aims	490
2	Methods	490
2.1	Data sources	490
2.2	Measures	491
2.3	Data analysis	493
3	Results	494
3.1	Descriptive statistics	494
3.2	Multi-ethnic participants' selection of a main ethnic group	495
3.3	Alignment between administrative-prioritisation and self-prioritisation	496
3.4	Effect of ethnic classification method on ethnic group proportions	497
4	Discussion	498
4.1	Validity of ethnicity prioritisation methods	499
4.2	Consistency of ethnic classification methods	500
4.3	Limitations	501
4.4	Implications	502
5	Conclusion	503
6	Acknowledgements	504
	References	505
	Appendix	509

Classifying multiple ethnic identifications: Methodological effects on child, adolescent, and adult ethnic distributions

Esther S. Yao¹

Kane Meissel²

Pat Bullen²

Polly Atatoa Carr³

Terryann C. Clark⁴

Susan M.B. Morton⁵

Abstract

BACKGROUND

The burgeoning global multi-ethnic population, in conjunction with the importance of accurate ethnic group counts for research and policy purposes, make classification of multiple ethnic responses a complex but important issue. There are numerous possible classification approaches, differing in ethical implications and ease of statistical application.

OBJECTIVE

This study empirically examines the validity and consistency of three comparatively accessible ethnic classification methods (total response, administrative-prioritisation, and self-prioritisation) in increasingly ethnically diverse age cohorts (adults, adolescents, and children).

METHOD

We utilised secondary data from two large-scale studies in Aotearoa/New Zealand which asked children ($N = 6,149$; responded via mother proxy), adolescents ($N = 8,464$), and adults ($N = 11,210$) to select (1) all the ethnicities they identified with, and (2) their main

¹ Faculty of Education and Social Work, The University of Auckland, Private Bag 92019, Victoria Street West, Auckland 1142, New Zealand. Email: e.yao@auckland.ac.nz.

² Faculty of Education and Social Work, The University of Auckland, New Zealand.

³ National Institute of Demographic and Economic Analysis, The University of Waikato, New Zealand.

⁴ School of Nursing, The University of Auckland, New Zealand.

⁵ Centre for Longitudinal Research, The University of Auckland, New Zealand.

ethnicity. The data were coded, then analysed using descriptive statistics and *z*-tests for proportional differences.

RESULTS

The majority of multi-ethnic participants were able to select a main ethnic group when required, but around 20% could not or refused to do so, and there was over 60% discrepancy between self-prioritised ethnicity and administrative-prioritised ethnicity. Differences by age group and ethnic combination were apparent. Comparison of overall ethnic group proportions outputted by the three classification methods revealed within-group variation, particularly where there were higher rates of multi-ethnic identification.

CONTRIBUTION

This study empirically demonstrates that researchers' choice of ethnic classification method can have a strong influence on ethnic group proportions. Researchers should therefore select the classification method most appropriate for their research question and clearly report the method employed.

1. Introduction

Ethnicity is a variable widely used to measure and analyse differences between population subgroups for research and policy purposes (Balestra and 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 2018; Perez and Hirschman 2009).⁶ 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 in 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 and Deane 2010; Didham 2005; Herman 2011). However, there is scant research that comprehensively investigates the effect 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 and 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

⁶ For brevity, 'multiple' refers to 'two or more' in this paper. The word 'dual' will be used when specifically referring to two ethnic groups.

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).

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).⁷ 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 2018; Bhopal 2004; Morning 2008).

Globally, a steadily growing proportion of the population report identification with more than one ethnic group (Aspinall 2018; Balestra and Fleischer 2018). In Aotearoa/New Zealand, 2018 census data showed that 11.4% of the population reported multiple ethnicities, up from 9.0% in 2001, 10.4% in 2006, and 11.2% in 2008 (Statistics New Zealand 2014, 2020). 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 (Jones and 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 group, compared to just 2.8% of those aged 65 years and over (Statistics New Zealand 2020). 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 2018; Perez and Hirschman 2009).

Recognising the increasing global trend in multiple ethnic identifications, the United Nations (2008: 139) updated the Principles and Recommendations for Population and Housing Censuses to require that “respondents have the option of indicating multiple

⁷ The term ‘ethnicity’, rather than ‘race’, will be used throughout this paper except when it pertains directly to the US context. In these instances, the use of race is synonymous with ethnicity.

ethnic affiliations” 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 and Fleischer 2018; United Nations 2008). However, common data collection methods for those identifying with multiple ethnicities can be classified into four categories (Aspinall 2018; 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.

1.2 Ethnic classification methods

Ethnicity data are commonly used to research and monitor ethnic inequities in health and social outcomes (Balestra and 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 anti-discrimination 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 and Deane 2010; Subramanian 2009), academic journals (Atatoa Carr et al. 2017; Herman 2011), and reports (Didham 2005; 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 1 provides an overview of common ethnic classification

methods.⁸ 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 typically not made transparent in academic journal publications, so patterns of usage are not readily known.

Table 1: Overview of common ethnic classification methods

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

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

⁸ 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).

contexts, particularly for longitudinal analyses or where numerator and denominator data are collected at different times or in different 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).

1.2.1 Retention of multiple ethnicities

Classification methods that retain multiple ethnicity data preserve multiple ethnic affiliations when the data are tabulated. Their main advantage is the congruence with respondent affiliation. However, these methods tend to pose more statistical challenges. Single/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). Single/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; and Middle Eastern, Latin American, or African (MELAA). Therefore, single/combination grouping would comprise European only, Māori only, Pacific only, European/Māori, European/Pacific, European/Māori/Pacific, 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).

1.2.2 Reduction of multiple ethnicities

By 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; now Statistics New Zealand). 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),⁹ 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-

⁹ 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.

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 and 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; Harris and Sim 2002) and is rarely observed in population 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 and 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, Markus, and Bergsieker 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 (Harris and Sim 2002).

1.3 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.

1.3.1 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 2020). 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, Boddington, and Didham 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, Boddington, and Didham 2000). In the 2018 Census, 11% of the population reported multiple ethnicities (Statistics New Zealand 2020). 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 2020). 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 single/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: 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).

1.3.2 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 administrative-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, administrative-prioritised, and self-prioritised output?

2. Method

2.1 Data sources

The current study utilises data from two large studies in Aotearoa/New Zealand: Growing Up in New Zealand (GUiNZ) and the Youth2000 National Youth Health and Wellbeing Survey. GUiNZ 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 GUiNZ 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 GUiNZ was obtained from the Ministry of Health Northern Y Regional Ethics Committee. All GUiNZ 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.

2.2 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. GUiNZ 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 GUiNZ survey also allowed participants to refuse response. Table 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 2) but were aggregated into 5 broad groupings for the purposes of this study: European,¹⁰ 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

¹⁰ 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).

ethnic classification methods used in this study were derived from aggregated total response data. Specifically, single/combination ethnicity was coded from total response ethnicity into 13 groups (see Table 3 for a group list). Finer groupings were not possible due to small cell counts ($n < 20$). Administrative-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.

Table 2: Ethnicity questions and response options by age group

	Children ^a	Adolescents	Adults
<i>Dataset</i>			
Study	Growing Up in New Zealand	Youth2000 survey series	Growing Up in New Zealand
Data collection wave	54 months	Youth'12	Antenatal
<i>Total response ethnicity</i>			
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")
<i>Self-prioritised ethnicity</i>			
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 <u>main</u> 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

Notes ^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., Samoan, 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).

A binary variable was then created to indicate instances where there was a discrepancy between administrative-prioritised ethnicity and self-prioritised ethnicity for multi-ethnic participants. In this study a discrepancy was operationalised as a case where administrative-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.

2.3 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 administrative-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, Miles, and Field 2012), an alpha level of $\alpha = .001$ was adopted to indicate statistical significance.¹¹

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 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 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

¹¹ Statisticians increasingly deem dichotomisation of statistical significance problematic (see Wasserstein, Schirm, and Lazar 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 the Appendix, and statistical testing was complemented by effect-size calculations.

weights were not required for child and adult data as the GUINZ 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).

3. Results

3.1 Descriptive statistics

Participants' ethnic identification (which, for children, was selected via mother proxy) are presented by age group using single/combination grouping in Table 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 3 shows that although the majority of participants in each age group identified with one broad ethnic grouping, there was higher ethnic diversity among the 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 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%).

Table 3: Single/combination ethnic identification by age group

	Children ^a		Adolescents		Adults	
	<i>n</i>	Col. %	<i>n</i>	Col. %	<i>n</i>	Col. %
<i>N</i>	6,149		8,464		11,210	
<i>1 ethnic group</i>						
European only	2221	36%	3989	47%	5639	50%
Māori only	168	3%	295	3%	671	6%
Pacific only	531	9%	546	6%	1256	11%
Asian only	604	10%	760	9%	1543	14%
Other only	266	4%	162	2%	343	3%
Total	3790	62%	5752	68%	9452	84%
<i>2 ethnic groups</i>						
European/Māori	725	12%	986	12%	974	9%
European/Pacific	195	3%	393	5%	217	2%
European/Asian	161	3%	235	3%	61	1%
European/Other	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	1846	30%	2262	27%	1596	14%
<i>3+ ethnic groups</i>						
European/Māori/Pacific	178	3%	126	1%	86	1%
Other 3+ ethnic groups	335	5%	324	4%	76	1%
Total	513	8%	450	5%	162	1%

Notes: ^a Child responses were collected by proxy from the child's mother.

Single/combination ethnicity information in Table 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’. European/Māori 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., European/Pacific, European/Māori/Pacific) 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.

3.2 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, a clear majority in each age group selected one main ethnic group.

Table 4: Dual- and multi-ethnic participants’ self-selection of a main ethnicity age group

	Children ^a		Adolescents		Adults	
	<i>n</i>	Col. %	<i>n</i>	Col. %	<i>n</i>	Col. %
<i>N</i>	2,359		2,712		1,758	
<i>Selected one main ethnic group</i>						
Matched admin-prioritisation	732	31%	1118	41%	677	39%
Did not match admin-prioritisation	1082	46%	1483	55%	733	42%
Total	1814	77%	2601	96%	1410	80%
<i>Did not select one main ethnic group ^b</i>						
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%

Notes. 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.

Pairwise age group comparisons conducted using two-sample z -tests (see Table A-1 in the Appendix) 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 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 Table A-2 in the Appendix. 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.

3.3 Alignment between administrative-prioritisation and self-prioritisation

Administrative-prioritisation is another method that can be used to categorise multi-ethnic identifications in a small number of mutually exclusive ethnic groupings. Our second research question investigated the rate of alignment between administrative-prioritised and self-prioritised ethnicity for multi-ethnic participants in each age group. Frequencies of instances where administrative-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. 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 Table A-3 in the Appendix) 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 Table A-2 in the Appendix. Once again, there was considerable within-group heterogeneity. For example, the discrepancy rate was higher in European/Asian children (92%), European/Asian adults (79%), and European/Other adolescents (85%) when compared to their European/Māori counterparts (children = 76%; adolescents = 65%; adults = 63%).

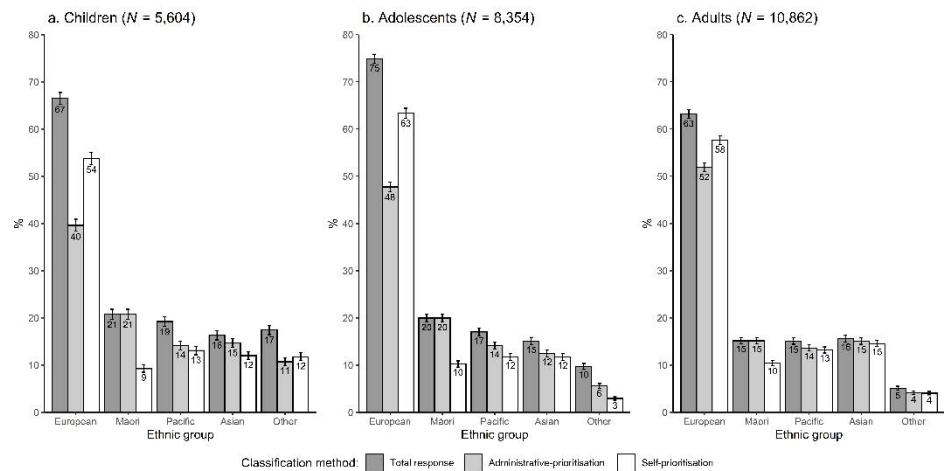
3.4 Effect of ethnic classification method on ethnic group proportions

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) 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 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 Tables A-4 to A-6 in the Appendix) 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 (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 Tables A-4 and A-5 in the Appendix), 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 Table A-6 in the Appendix), 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 $p > .001$).

Figure 1: Ethnic group proportions by ethnic classification method and age group



Notes: 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 (CIs).

4. 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 2018; Statistics New Zealand 2014), 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 administrative-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.

4.1 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 and 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 (Harris and 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 (Khawaja, Boddington, and Didham 2000).

The second research question examined the rate of alignment between self-prioritised and administrative-prioritised ethnicity for multi-ethnic participants in each age group. Consistent with previous studies (e.g., Atatoa Carr et al. 2017; Kukutai and 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 GUINZ 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 administrative-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 administrative-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, as 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 need to be assessed in light of each research or policy issue under investigation.

4.2 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 administrative-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 and 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 and Hirschman 2009). This supports the finding from 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.

4.3 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 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 (Harris and Sim 2002; Liebler et al. 2017), so ethnicity responses cannot be interpreted as a static characteristic of participants.

4.4 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 and 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 2018; Balestra and Fleischer 2018; Perez and Hirschman 2009; Statistics New Zealand 2014), 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 single/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 (Harris and Sim 2002; Herman 2004); and asking for a self-prioritised ethnicity may be ethically contentious or offensive to some participants (Aspinall and Song 2013; Sanchez 2010; Townsend, Markus, and Bergsieker 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, single/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 ethnic data that allow the most suitable classification method to be used (Aspinall 2018).

5. Conclusion

The significant effect of ethnic classification methods 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.

6. Acknowledgements

We would like to acknowledge the participants of the GUiNZ and Youth'12 studies, as well as the members of the GUiNZ team and the Adolescent Health Research Group (AHRG), for making this study possible. Special thanks to Jemaima Tiatia-Seath (University of Auckland) for her helpful comments on an earlier version of this manuscript. The views expressed in this paper are those of the authors and do not necessarily represent the views of GUiNZ or AHRG.

References

- Aspinall, P.J. (2018). 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. doi:10.1080/01419870.2017.1346267.
- Aspinall, P.J. and Song, M. (2013). *Mixed race identities*. London: Palgrave Macmillan. doi:10.1057/9781137318893.
- Atatoa Carr, P.E., Bandara, D., Berry, S., Kingi, T.K., Grant, C., and Morton, S. (2017). Ethnic identification complexity across generations: Evidence from growing up in New Zealand. *New Zealand Population Review* 43: 35–61.
- Balestra, C. and Fleischer, L. (2018). *Diversity statistics in the OECD: How do OECD countries collect data on ethnic, racial and indigenous identity?* Paris: OECD Publishing (OECD Statistics Working Papers 2018/09).
- Bhopal, R. (2004). Glossary of terms relating to ethnicity and race: For reflection and debate. *Journal of Epidemiology and Community Health* 58(6): 441–445. doi:10.1136/jech.2003.013466.
- Boven, N., Exeter, D., Sporle, A., and 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. doi:10.1080/1177083X.2019.1657912.
- Callister, P., Didham, R., Potter, D., and Blakely, T. (2007). Measuring ethnicity in New Zealand: Developing tools for health outcomes analysis. *Ethnicity and Health* 12(4): 299–320. doi:10.1080/13557850701300699.
- 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., and The Adolescent Health Research Group. (2013). *Youth'12 prevalence tables: The health and wellbeing of New Zealand secondary school students in 2012*. Auckland: The University of Auckland.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: 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. doi:10.1037/0022-0167.54.3.224.

- Denton, N.A. and Deane, G.D. (2010). Researching race and ethnicity: Methodological issues. In: Collins, P.H. and Solomos, J. (eds.). *The SAGE handbook of race and ethnic studies*. London: Sage: 67–89. doi:10.4135/9781446200902.n5.
- Department of Statistics (1993). *New Zealand standard classification of ethnicity*. Wellington: Author.
- Didham, R. (2005). *Understanding and working with ethnicity data*. Wellington: Statistics New Zealand.
- Education Counts (2014). Ethnic group codes. Wellington: Education Counts. https://www.educationcounts.govt.nz/data-services/collecting-information/code-sets-and-classifications/ethnic_group_codes.
- Field, A., Miles, J., and Field, Z. (2012). *Discovering statistics using R*. London: Sage.
- Harris, D.R. and Sim, J.J. (2002). Who is multiracial? Assessing the complexity of lived race. *American Sociological Review* 67(4): 614–627. doi:10.2307/3088948.
- Herman, M. (2004). Forced to choose: Some determinants of racial identification in multiracial adolescents. *Child Development* 75(3): 730–748. doi: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. doi:10.1111/j.1751-9020.2011.00388.x.
- Ingram, D.D., Parker, J.D., Schenker, N., Weed, J.A., Hamilton, B., Arias, E., and Madans, J.H. (2003). United States Census 2000 population with bridged race categories. (Vital and Health Statistics 2/135). Washington, D.C.: National Center for Health Statistics.
- Jones, N.A. and Bullock, J. (2012). The two or more races population: 2010. Washington, D.C.: U.S. Census Bureau (2010 Census Brief Series).
- Khawaja, M., Boddington, B., and Didham, R. (2000). *Growing ethnic diversity in New Zealand and its implications for measuring differentials in fertility and mortality*. Wellington: Statistics New Zealand.
- Kukutai, T. and Callister, P. (2009). A ‘main’ ethnic group? Ethnic self-prioritisation among New Zealand youth. *Social Policy Journal of New Zealand* 36: 16–31.
- Lee, S.M. (2001). Using the new racial categories in the 2000 Census. (KIDS COUNT/PRB Reports on Census 2000). Baltimore: Annie E. Casey Foundation.

- Liebler, C.A., Porter, S.R., Fernandez, L.E., Noon, J.M., and 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. doi:10.1007/s13524-016-0544-0.
- Mays, V.M., Ponce, N.A., Washington, D.L., and Cochran, S.D. (2003). Classification of race and ethnicity: Implications for public health. *Annual Review of Public Health* 24(1): 83–110. doi:10.1146/annurev.publhealth.24.100901.140927.
- Ministry of Health (2017). *HISO 10001:2017 Ethnicity Data Protocols*. Wellington: Ministry of Health.
- 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. doi: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., and Wall, C. (2013). Cohort profile: Growing up in New Zealand. *International Journal of Epidemiology* 42(1): 65–75. doi:10.1093/ije/dyr206.
- Morton, S.M.B., Ramke, J., Kinloch, J., Grant, C.C., Atatoa Carr, P., Leeson, H., Lee, A.C.L., and 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. doi:10.1111/1753-6405.12220.
- 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*. Washington, D.C.: Executive Office of the President.
- Perez, A.D. (2006). *Choosing a race: The effects of proxy reporting on the identification of mixed race children*. Paper presented at the Annual Meeting of the Population Association of America, Los Angeles, CA, March 30–April 1, 2006.
- Perez, A.D. and Hirschman, C. (2009). The changing racial and ethnic composition of the US population: Emerging American identities. *Population and Development Review* 35(1): 1–51. doi:10.1111/j.1728-4457.2009.00260.x.
- R Core Team (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org>.

- Rutkowski, R.E., Salemi, J.L., Tanner, J.P., Matas, J.L., and 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.
- 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. doi:10.1111/j.1559-1816.2010.00634.x.
- Statistics New Zealand (2004). *Report of the review of the measurement of ethnicity*. Wellington: Statistics New Zealand.
- Statistics New Zealand (2005). *Statistical standard for ethnicity*. Wellington: Statistics New Zealand.
- Statistics New Zealand (2014). *2013 Census QuickStats about culture and identity*. Wellington: Statistics New Zealand.
- Statistics New Zealand (2020). 2018 Census ethnic group summaries. Wellington: Statistics New Zealand. <https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries>.
- Subramanian, S. (2009). Methods and approaches in using secondary data sources to study race and ethnicity factors. In: Verma, M. (ed.). *Cancer epidemiology* (Vol. 1). Totowa: Humana Press: 227–237. doi:10.1007/978-1-59745-416-2_12.
- Townsend, S.S.M., Markus, H.R., and Bergsieker, H.B. (2009). My choice, your categories: The denial of multiracial identities. *Journal of Social Issues* 65(1): 185–204. doi:10.1111/j.1540-4560.2008.01594.x.
- United Nations (2008). *Principles and recommendations for population and housing censuses* (Revision 2). New York: United Nations.
- Wasserstein, R.L., Schirm, A.L., and Lazar, N.A. (2019). Moving to a world beyond ‘ $p < 0.05$ ’. *American Statistician* 73(Supplement 1): 1–19. doi:10.1080/00031305.2019.1583913.

Appendix

Table A-1: Two-sample z-tests of pairwise age differences in dual- and multi-ethnic participants' selection rate of one main ethnic group

Age group	Group A			Group B			Group A vs. Group B				
	<i>N</i>	<i>n</i>	Row %	Age group	<i>N</i>	<i>n</i>	Row %	% Difference (A-B)	<i>z</i>	<i>p</i>	<i>ES</i> ^a
Children ^b	2,359	1,814	77%	Adolescents	2,712	2,601	96%	-19%	-20.12	<.001	-0.60
Children ^b	2,359	1,814	77%	Adults	1,758	1,410	80%	-3%	-2.55	.011	-0.08
Adolescents	2,712	2,601	96%	Adults	1,758	1,410	80%	16%	16.89	<.001	0.51

Notes: ^a Cohen's (1988) *h* effect size.

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

Table A-2: Dual- and multi-ethnic participants' self-selection of a main ethnic group by age group and ethnic combination

	<i>N</i>	Selected one main ethnic group				Did not select one main ethnic group	
		Matched admin-prioritisation		Did not match admin-prioritisation		<i>n</i>	Row %
		<i>n</i>	Row %	<i>n</i>	Row %		
<i>Children</i> ^a (<i>N</i> = 2,359)							
European/Māori	725	167	23%	366	50%	192	26%
European/Pacific	195	69	35%	84	43%	42	22%
European/Asian	161	13	8%	100	62%	48	30%
European/Other	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%
European/Māori/Pacific	178	58	33%	70	39%	50	28%
Other 3+ ethnic groups	335	61	18%	204	61%	70	21%
<i>Adolescents</i> (<i>N</i> = 2,712)							
European/Māori	987	350	35%	618	63%	19	2%
European/Pacific	393	225	57%	S	41%	<10	2%
European/Asian	234	100	43%	115	49%	19	8%
European/Other	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%
European/Māori/Pacific	126	S	48%	62	49%	<10	3%
Other 3+ ethnic groups	325	147	45%	156	48%	22	7%
<i>Adults</i> (<i>N</i> = 1,758)							
European/Māori	974	361	37%	412	42%	201	21%
European/Pacific	217	87	40%	78	36%	52	24%
European/Asian	61	S	21%	40	66%	<10	13%
European/Other	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%
European/Māori/Pacific	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.

Table A-3: Two-sample z-tests of age differences in the discrepancy rate between administrative-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	2,359	1,627	69%	Adolescents	2,712	1,594	59%	10%	7.52	<.001	0.21
Children ^b	2,359	1,627	69%	Adults	1,758	1,081	61%	7%	5.00	<.001	0.16
Adolescents	2,712	1,594	59%	Adults	1,758	1,081	61%	-3%	-1.81	.071	-0.06

Notes: ^a Cohen's (1988) *h* effect size.

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

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	
<i>European</i>										
Total response	3,727	67%	Self-prioritisation	3,016	54%	-13%	-20.12	<.001	-0.26	
Total response	3,727	67%	Admin-prioritisation	2,221	40%	-27%	-42.62	<.001	-0.55	
Self-prioritisation	3,016	54%	Admin-prioritisation	2,221	40%	-14%	-21.30	<.001	-0.29	
<i>Māori</i>										
Total response	1,165	21%	Self-prioritisation	519	9%	-12%	-21.27	<.001	-0.33	
Total response	1,165	21%	Admin-prioritisation	1,165	21%	-	-	-	-	
Self-prioritisation	519	9%	Admin-prioritisation	1,165	21%	12%	29.77	<.001	0.33	
<i>Pacific</i>										
Total response	1,077	19%	Self-prioritisation	734	13%	-6%	-11.63	<.001	-0.17	
Total response	1,077	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	
<i>Asian</i>										
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	
<i>Other</i>										
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	

Notes: ^a Child responses were collected by proxy from the child's mother.

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

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
<i>European</i>									
Total response	6,259	75%	Self-prioritisation	5,293	63%	-12%	-24.38	<.001	-0.25
Total response	6,259	75%	Admin-prioritisation	3,989	48%	-27%	-57.30	<.001	-0.57
Self-prioritisation	5,293	63%	Admin-prioritisation	3,989	48%	-16%	-29.61	<.001	-0.32
<i>Māori</i>									
Total response	1,669	20%	Self-prioritisation	855	10%	-10%	-22.27	<.001	-0.28
Total response	1,669	20%	Admin-prioritisation	1,669	20%	-	-	-	-
Self-prioritisation	855	10%	Admin-prioritisation	1,669	20%	10%	29.38	<.001	0.28
<i>Pacific</i>									
Total response	1,423	17%	Self-prioritisation	981	12%	-5%	-12.86	<.001	-0.15
Total response	1,423	17%	Admin-prioritisation	1,181	14%	-3%	-7.04	<.001	-0.08
Self-prioritisation	981	12%	Admin-prioritisation	1,181	14%	2%	6.80	<.001	0.07
<i>Asian</i>									
Total response	1,262	15%	Self-prioritisation	980	12%	-3%	-8.62	<.001	-0.10
Total response	1,262	15%	Admin-prioritisation	1,044	12%	-3%	-6.66	<.001	-0.08
Self-prioritisation	980	12%	Admin-prioritisation	1,044	12%	1%	2.18	.030	0.02
<i>Other</i>									
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

Notes: ^a Cohen's (1988) *h* effect size.**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
<i>European</i>									
Total response	6,865	63%	Self-prioritisation	6,266	58%	-6%	-11.92	<.001	-0.11
Total response	6,865	63%	Admin-prioritisation	5,639	52%	-11%	-24.39	<.001	-0.23
Self-prioritisation	6,266	58%	Admin-prioritisation	5,639	52%	-6%	-12.18	<.001	-0.12
<i>Māori</i>									
Total response	1,649	15%	Self-prioritisation	1,134	10%	-5%	-13.77	<.001	-0.14
Total response	1,649	15%	Admin-prioritisation	1,649	15%	-	-	-	-
Self-prioritisation	1,134	10%	Admin-prioritisation	1,649	15%	5%	16.16	<.001	0.14
<i>Pacific</i>									
Total response	1,639	15%	Self-prioritisation	1,436	13%	-2%	-5.44	<.001	-0.05
Total response	1,639	15%	Admin-prioritisation	1,481	14%	-1%	-4.24	<.001	-0.04
Self-prioritisation	1,436	13%	Admin-prioritisation	1,481	14%	0%	1.27	.202	0.01
<i>Asian</i>									
Total response	1,697	16%	Self-prioritisation	1,583	15%	-1%	-3.01	.003	-0.03
Total response	1,697	16%	Admin-prioritisation	1,638	15%	-1%	-1.56	.119	-0.02
Self-prioritisation	1,583	15%	Admin-prioritisation	1,638	15%	1%	1.50	.135	0.01
<i>Other</i>									
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

Notes: ^a Cohen's (1988) *h* effect size.

