

Table of Contents



Table of Contents 2

1. Executive Summary 6
2. Introduction 7
3. Overview of the Data 8
   1. Accounts in Scope - Survivors with Accounts 8
   2. Data Sources: What did DOT receive? 8
   3. Data Collection Quality 8
   4. Methodology 9
   5. Data Completeness 11
   6. Data Enhancement 11
   7. Confidentiality 14
4. Key Insights 14
5. Demographic Analysis 22
   1. Current Age of Survivors 22
   2. Current Age by Gender 23
   3. Ethnicity by Gender 24
   4. Current Age by Ethnicity 25
   5. Deaf or Disability by Age 26
   6. Deaf or Disability by Ethnicity 27
   7. Experience of Incarceration 28
      1. Incarceration by Gang Affiliation 28
      2. Incarceration by Ethnicity 29
      3. Incarceration by Gender 30
      4. Incarceration by Deaf and Disability 31
6. Pathways Into Care 33
   1. Findings for Pathways into Care 33
   2. Categories of Pathways Into Care 33
   3. Pathway into Care by Decade of Abuse 36
7. Abuse Types 38
   1. Relationship between Abuse Types 38
   2. Relationship between Abuse Types 39
   3. Decade of Abuse by Abuse Types 40
   4. Distribution of Abuse Types by Decade of Abuse 41
      1. Decade of Abuse by Ethnicity 42
      2. Abuse Type by Age 43
   5. Abuse Type by Ethnicity 44

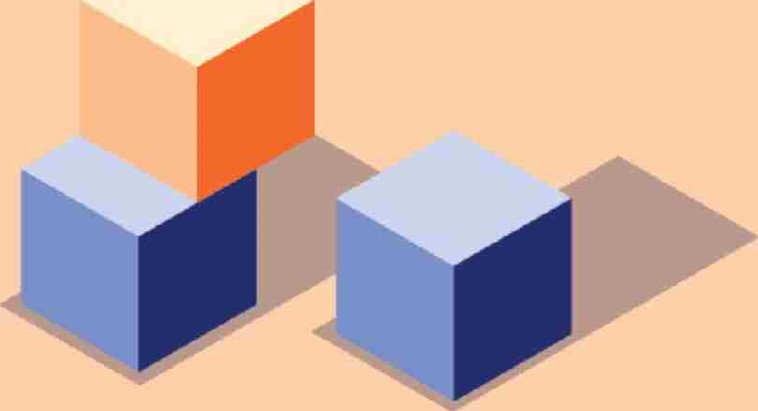
|  |  |
| --- | --- |
| 7.5.1. Proportions by Ethnicity | 45 |
| 7.5.1.1. Mäori more likely to report physical abuse | 45 |
| 7.5.1.2. Pasifika more likely to report physical abuse | 45 |
| 7.5.1.3. Other significant relationships between ethnicity and abuse type: | 46 |
| 7.6. Abuse Type by Gender | 46 |
| 7.6.1. Proportion of Abuse Type by Gender | 47 |
| 7.6.2. Males more likely to report Physical Abuse | 47 |
| 7.6.3. Females more likely to report Sexual Abuse | 47 |
| 7.6.4. Other Significant Relationships between Gender and Abuse Type: | 48 |
| 7.7. Abuse Type by Ethnicity and Gender | 48 |
| 7.7.1. Physical Abuse by Ethnicity and Gender | 48 |
| 7.7.2. Sexual Abuse by Ethnicity and Gender | 49 |
| 7.7.3. Emotional Abuse by Ethnicity and Gender | 50 |
| 7.7.4. Neglect by Ethnicity and Gender | 50 |
| 7.8. Abuse Type by Mental Distress | 51 |
| 7.9. Abuse Type by Sexual Identity | 51 |
| 7.10. Abuse Type by Deaf or Disability | 52 |
| 7.11. Abuse Type by Ethnicity and Deaf or Disability | 54 |
| 8. Institution and Setting | 56 |
| 8.1. Count by Institution | 56 |
| 8.2. Count by Setting | 56 |
| 8.3. Setting by Length of Time in Care | 58 |
| 8.4. Setting by Experience of Incarceration | 59 |
| 8.5. Count of Institutions by Incarceration | 61 |
| 8.6. Institution Types by Abuse Type | 62 |
| 8.7. Setting by Abuse Type | 62 |
| 8.8. Institution by Abuse Type | 64 |
| 8.9. Demographic Prevalence at Institutions and Settings | 66 |
| 8.9.1. Mental Distress | 66 |
| 8.9.2. Mental Distress by Institution type | 66 |
| 8.9.3. Mental Distress by Setting | 66 |
| 8.9.4. Deaf and Disability | 69 |
| 8.10. Abuse Type by Deaf or Disability | 70 |
| 8.10.1. Deaf or Disability Types by Setting Type | 71 |
| 8.10.2. Deaf or Disability by Institution Type | 72 |
| 9. Conclusion | 74 |
| 10. Appendix | 75 |
| 10.1. Definitions of Variables | 75 |
| 10.2. Definitions of Groupings within Variables | 77 |
| 10.3. Iwi Classification | 77 |
| 10.4. Excluding Accounts Out of Scope | 79 |

DOT

|  |  |
| --- | --- |
| 10.5. DOT's Data Security and Storing | 79 |
| 10.6. Methodology | 80 |
| 10.6.1. Data Loading and Preprocessing | 80 |
| 10.6.2. Text Processing and Tokenization | 80 |
| 10.6.3. Keyword Tagging | 80 |
| 10.6.4. Data Modelling | 80 |
| 10.6.5. Prediction Review | 81 |
| 10.6.6. Data Imputation | 81 |
| 10.7. Supplementary Tables | 82 |



# Overview



## 1. Executive Summary

The Abuse in Care Royal Commission of Inquiry is an exhaustive investigation into abuse suffered by tamariki, rangatahi and adults in state and faith-based care in Aotearoa New Zealand between the years 1950 and 1999. Accounts from 2,329 survivors, who registered with the Inquiry, form the heart of this report.

The work DOT Loves Data (DOT) has completed to date is an important component of the Royal Commission of Inquiry. It provides a quantitative analysis of survivor accounts. It also ensures a consistent and objective analysis of all survivor accounts. In the initial stages of the Royal Commission, Inquiry staff manually reviewed witness statements and transcripts of interviews, identifying patterns and employing quantitative analysis on coded data. DOT's involvement focused on quantitative analysis of survivor accounts using natural language processing and aimed to bridge any data gaps and bolster data integrity.

We were able to analyse factors such as survivor ethnicity, gender, LGBTQIA+ identity, mental distress and disabilities against the types of institution, care settings and abuse. Results from DOT's analysis quantified insights, including the prevalence and types of abuse reported in the survivor accounts and the personal characteristics or factors that might have influenced their care experience.

Of the survivors who spoke with the Inquiry, our analysis showed that the number of incidents of reported abuse peaked in the 1970s, coinciding with the period of highest population in state and faith-based care facilities. The most common pathways into care were a state requirement due to troublesome behaviour, voluntary placement by parents due to a lack of support or insufficient financial means, or admission into a faith-based school.

Some survivor cohorts reported especially high levels of abuse, with a disproportionately high representation of Mäori, Deaf, survivors living with a disability and Pasifika. Our analysis showed that Mäori experienced higher care admission rates relative to the general population (Mäori make up 44% of survivor accounts), while Mäori and Pasifika experienced higher levels of physical abuse. Similarly, survivors who are Deaf or living with a disability experienced elevated abuse rates compared to those who are not Deaf or living without a disability.

The age and gender of survivors also played a significant role in the abuse they were subjected to. Survivors aged 10 - 14, for example, reported the highest levels of sexual and physical abuse by their caregivers. In terms of gender, 59% of survivor accounts were male, 40% female and 1% gender diverse, non-binary or other. Male survivors reported higher levels of physical and solitary abuse, while female survivors reported higher levels of sexual, emotional, neglect and medical abuse.

Survivors' life outcomes were significantly and detrimentally impacted by their abuse in care, relative to the general population. 83% of survivors reported suffering mental distress challenges during or following their time in care. 29% or survivors had been imprisoned and 10% of survivors had joined a gang.

The repercussions of the abuse and neglect suffered by survivors are evident within the analysis. In the pages that follow, we will analyse the survivor accounts to quantify and explain, as best as the data allows, why people were taken into care, what abuse took place and the life-long impacts of that abuse on survivors.

## 2. Introduction

The Abuse in Care Royal Commission of Inquiry is investigating abuse against tamariki, rangatahi and adults in care between 1950-1999. The Royal Commission has collected information from 2,329 survivors with accounts of abuse, noting that the data also takes into account survivor reports of their experience in care after 1999. We further note that our analysis relates only to these 2,329 survivor accounts and are conscious that there will be survivors who did not speak to the inquiry and for whom we hold no information.

In the initial stages of the Royal Commission, insight from the data was found by manually reading witness statements and private interviews by researchers searching for themes and patterns, and through quantitative analysis of coded data.

DOT is one of New Zealand's leading providers of data science, data analysis and data visualisation. We employ a range of PhD level researchers and data scientists to ensure we are able to solve and distil complex problems into actionable insights. DOT was engaged to provide a quantitative analysis of the data collected from the survivor accounts as defined by its terms of reference. Using natural language processing, our role was to assist the Royal Commission to understand any existing data gaps, to improve the data integrity, and explain the data gaps in an easy to understand way. As the data collection techniques were improved throughout the collection of the survivor accounts, DOT's remit was to ensure consistent analysis was captured across all accounts, based on both written and audio transcripts, in relation to features such as ethnicity, iwi, gender, LGBTQIA+, mental distress and disability. We also ensured that every survivor account was considered in detail and objectively.

3. Overview of the Data

## 3.1. Accounts in Scope - Survivors with Accounts

This analysis includes the experiences of 2,329 survivors with accounts. These survivors included in the analysis met the following criteria:

* They are survivors of abuse
* They are people with accounts that registered with the abuse in care Royal Commision (individuals for whom the Abuse in Care Royal Commission holds personal information with some detail on their abuse while in care). The accounts include survivor interviews, private sessions and written accounts - throughout the report this is referred to as asurvivor accounts".

Based on guidance from the Abuse in Care Royal Commission, we have only included survivors with accounts in this analysis. Throughout this report, these survivors with accounts are referred to as 'survivors'.

3.2. Data Sources: What did DOT receive?

We received an initial 200 survivor accounts as a sample data set so we could begin to analyse for distributions of missing information, scope audio and text data for size, complexity and completeness, prototype simple methods for extracting missing information and review our original estimates of the project timings. In total, we received 4,998 survivor-related documents from the Royal Commission, which consisted of witness statements, written or audio transcripts and other supporting documents that were filtered to include survivors with accounts.

We used our original sample analysis to build out our analysis against the 2/329 survivors with accounts supplied by the Royal Commission.

## 3.3. Data Collection Quality

We have highlighted some of the quality issues in the data and the limitations of this report due to the gaps in the data and data quality. Of the survivor accounts collected from survivors of abuse in care, there is a variety of data captured on both the demographics of the survivor and also the nature of the abuse endured.

The witness statements were collected by a range of interviewers and as time went on there was increased identification of the required information to be captured on the survivor by the interviewer. Due to this, the information collected from interviewer to interviewer varied greatly. Over time, the information collected also increased meaning that earlier witness statements contained less information than the more recent witness statements.

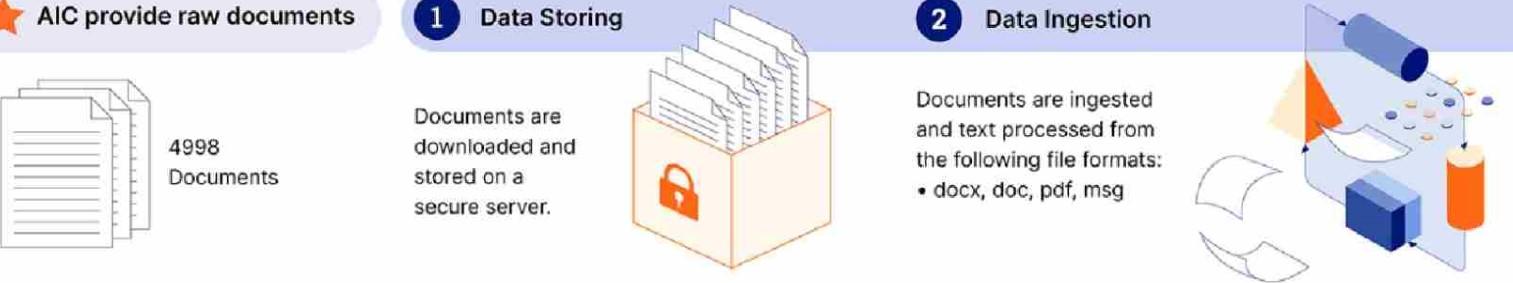
To combat the missing data and gaps in information required, we used natural language processing (NLP) to extract information from the free text fields, to fill in missing data fields and populate the dataset. Further information on this process can be found in detail in the methodology section.

### 3.4. Methodology

An overview of the process that we undertook is shown in the figure below:

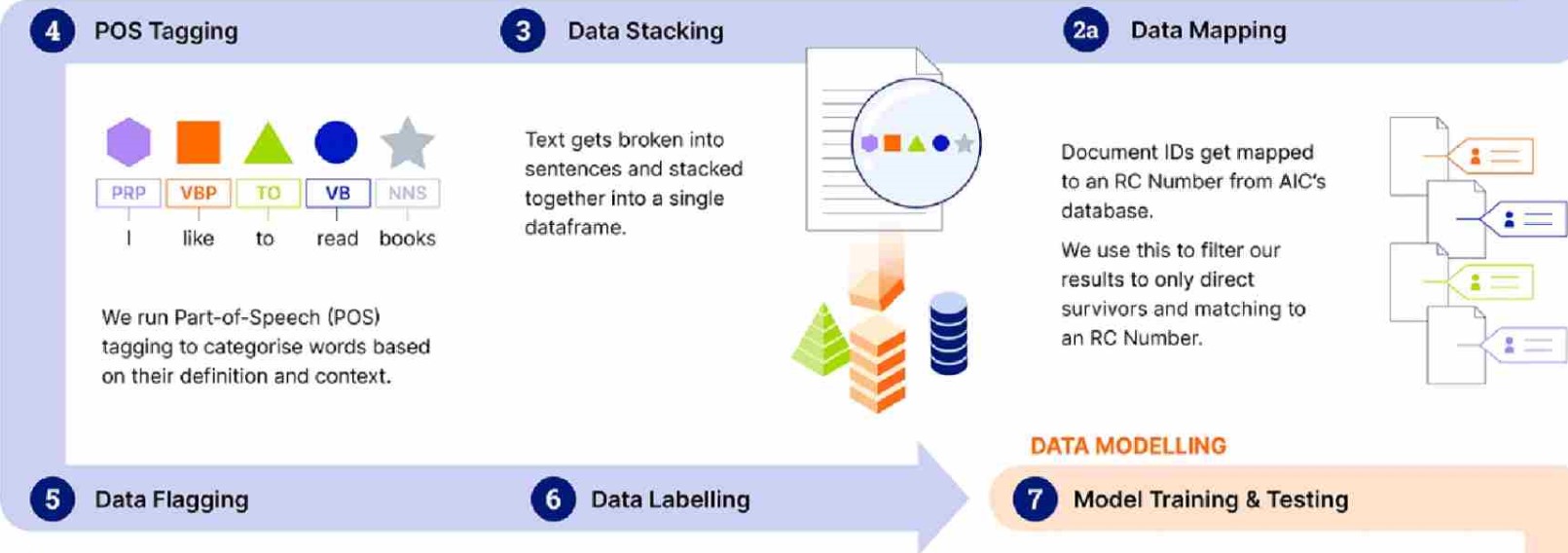
 How we worked e DOT AIC

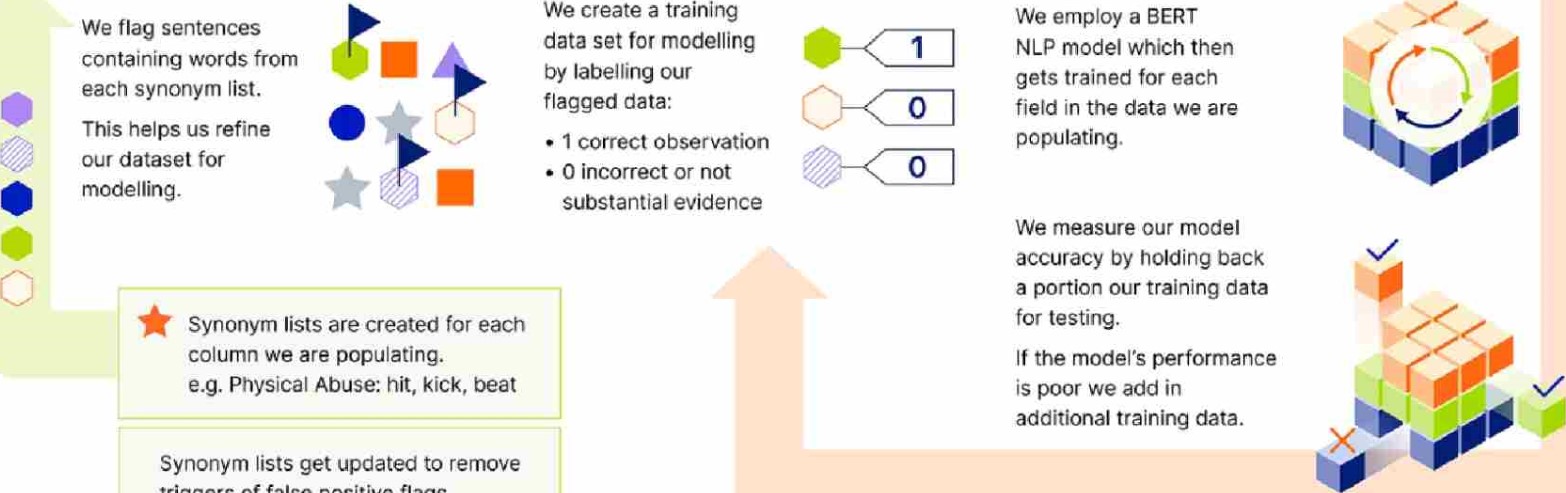
DATA ORGANISING



O

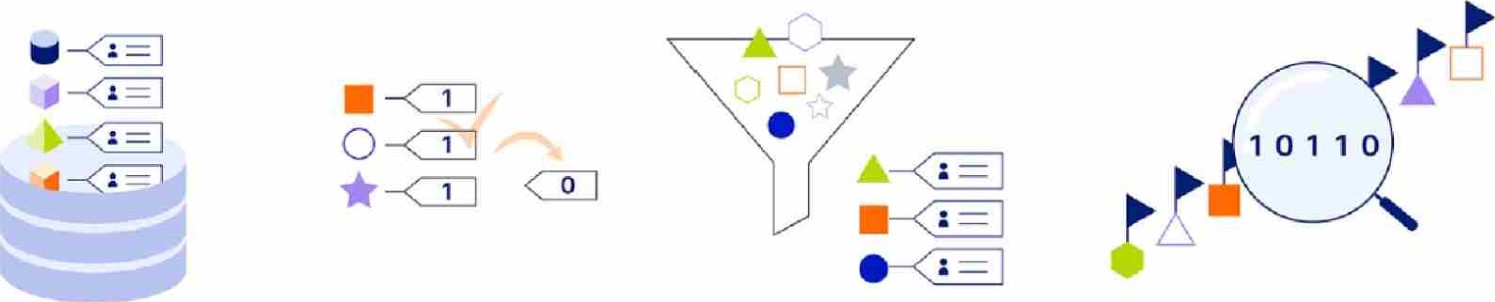
O



Synonym lists get updated to remove triggers of false positive flags.

FILTER & REVIEW

11 Final  Review O Filter & Collate by RC O Model Prediction



|  |  |  |  |
| --- | --- | --- | --- |
| Fields are combined and all | We review a subset of the | We filter our data to positive | Our models get run across the |
| positive observations are | positive observations for each | observations. For certain fields | flagged data to create our |
| related back to an RC | data field. Correcting false | we collate positive observations | prediction data set. |
| Number in our final dataset. | positives as necessary. | by an RC Number. |  |

Figure 1: Methodology - How We Worked

•

## 3.5. Data Completeness

Prior to integrating the enriched data into the existing Customer Relationship Management (CRM) system, DOT carried out a thorough and robust review of the new data variables produced. We manually reviewed 100% of 26 out of 29 attributes and undertook a random sample of between 25-50% for the remaining four attributes. These were: Physical Abuse, Sexual Abuse, Emotional Abuse and Institution.

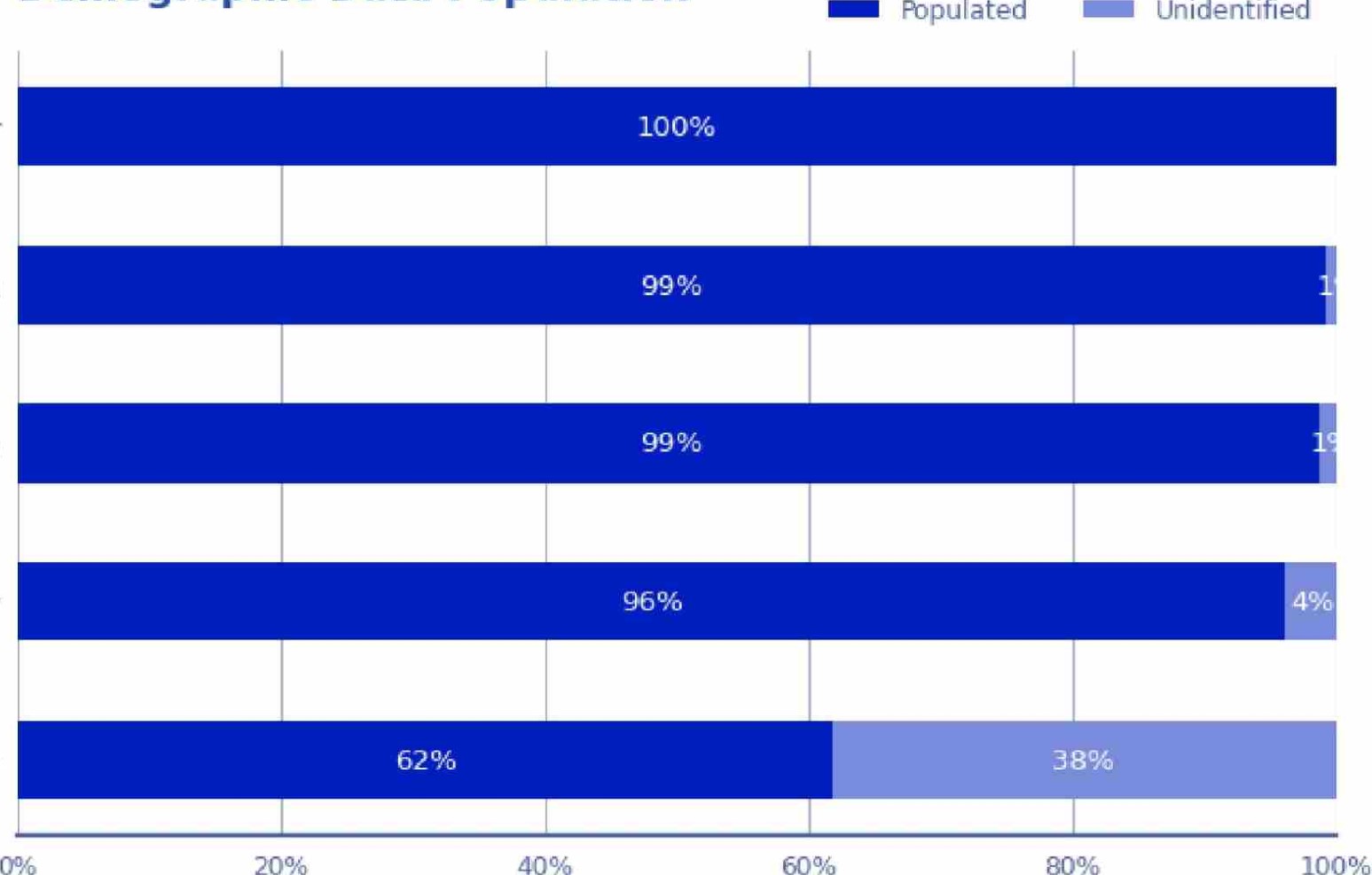
We raised with AIC that/ due to very tight timeframes, DOT was not able to manually check 100% of the positive cases for these four attributes. The agreed standard with AIC was a precision of 90% where precision is the true positive rate. This was achieved in the samples based on independent review by AIC.

DOT reviewed this with caution and took a conservative approach to ensure only true positives were counted and therefore integrated into the CRM. This ultimately means there will be undercounting due to the nature and quality of the data, which ensures that the counts noted throughout this report are baseline minimums.

### 3.6. Data Enhancement

In the figure below, we show the population of demographic variables in the final data set. These fields were largely complete prior to this work but if absent and adequate evidence was available these were further populated by DOT.

#### Demographic Data Population

Gender

Region

3

o Year of Birth

Ethnicity

LGBTQIA+

Proportion of Direct Survivors

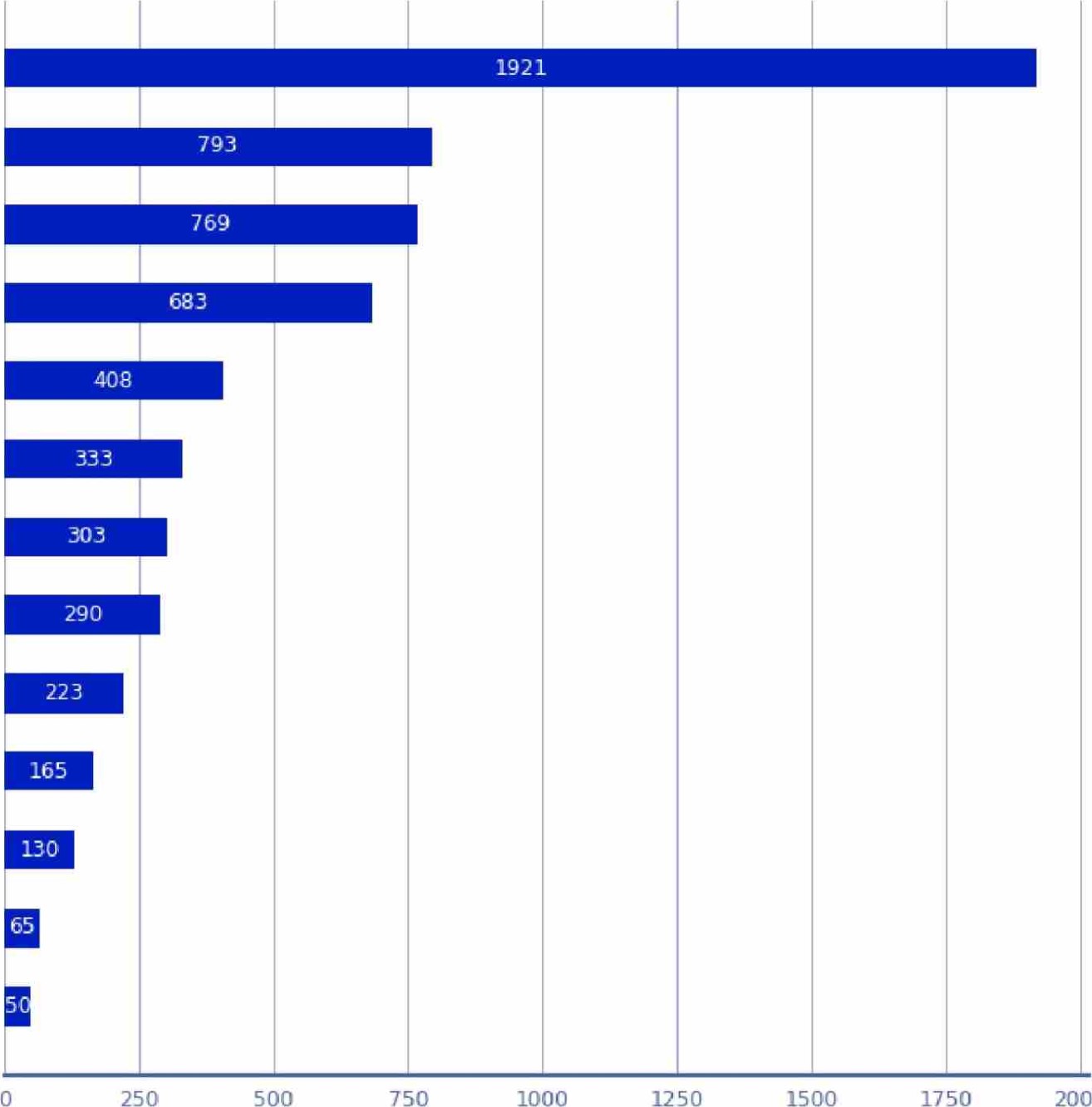
Figure 2: Demographic Data Population

Based on the work undertaken by the Abuse In Care Royal Commission of Inquiry and DOT's review, we have been able to capture survivor demographic variable data for gender (100%), region currently lived in (99%), year of birth (99%), Ethnicity (96%) and members of the LGBTQIA+ community (62%).

We assessed these for their completeness and accuracy. The graph below depicts the number of survivor counts by field. For the fields shown, enhancements were implemented only when positive observations of the variable were found; otherwise, they were left \_blank. For instance, the Disability - Blind' field is populated solely by people who are blind, rather than having a yes/no option. Our approach follows a conservative methodology, which might result in undercounting. However, discerning whether these values represent true or false negatives is challenging.

##### Survivor Count per Category

Mental Distress Conditions



2000

Iwi

Chronic Health Conditions

Ever Incarcerated

Self Harm

Gang Affiliation

Neurodivergent

Addiction

Disability - Learning

Disability - Mobility

Deaf

Disability - Blind

Disability - Communication

Number of Direct Survivors

Figure 3: Survivor Count per Category

The below graph shows the abuse data that was collected through this process. Due to this enrichment of the dataset, we are now able to carry out a quantitative analysis with certainty and accuracy when describing the nature and extent of abuse in care throughout this report.

#### Survivor Count per Abuse Category

Any Abuse

Physical

Sexual

Emotional

Neglect

Medical

Solitary

Non-Contact Sexual

0 2000

500 1000 1500

Number of Direct Survivors

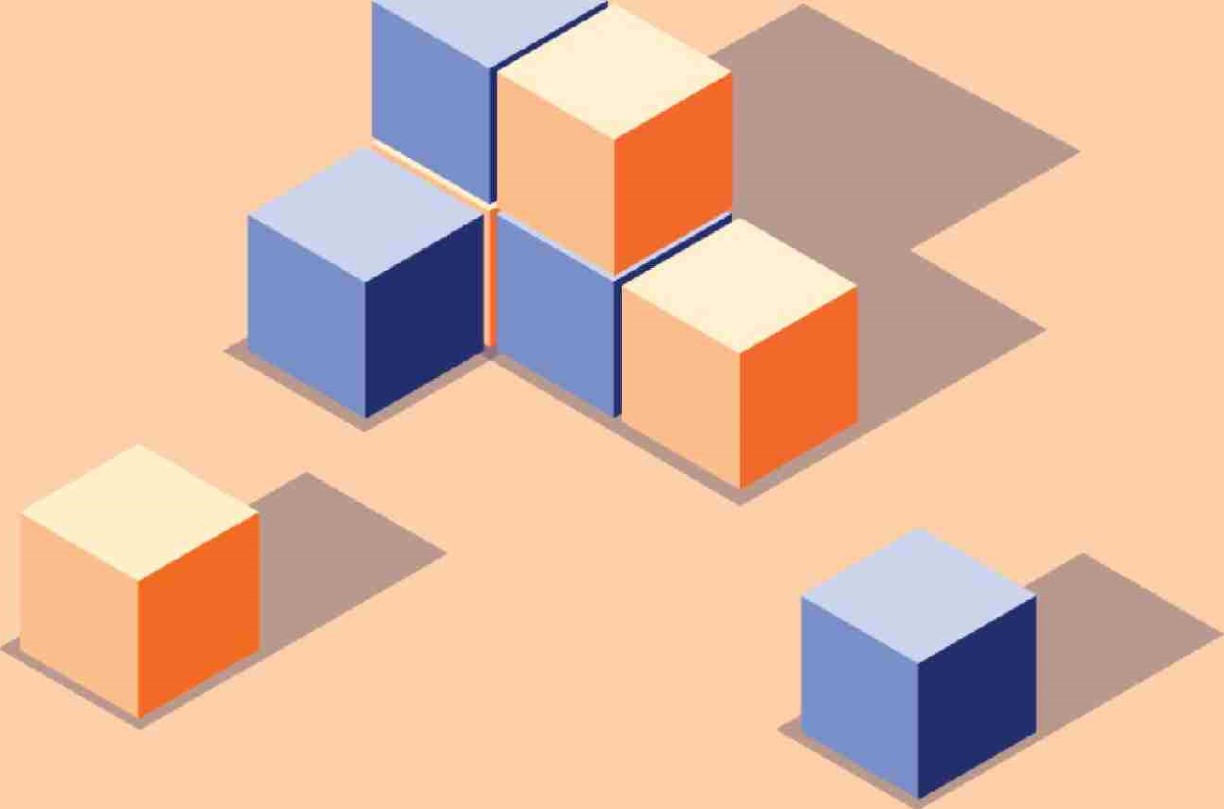
Figure 4: Survivor Count per Abuse Category

### 3.7. Confidentiality

For confidentiality purposes, some sections or variables may remove certain ethnicities or genders for the purpose of ensuring the confidentiality of the survivors. For example, where 'gender diverse, non-binary or other' count is fewer than six, it has been suppressed for confidentiality. For some analysis on ethnicity, the Asian, MEL AA and Other ethnicity groups have been grouped so they can be reported on. This has been done to maintain confidentiality where the individual type counts were too low, '..C' has been used in replacement for confidentiality of the actual figures .

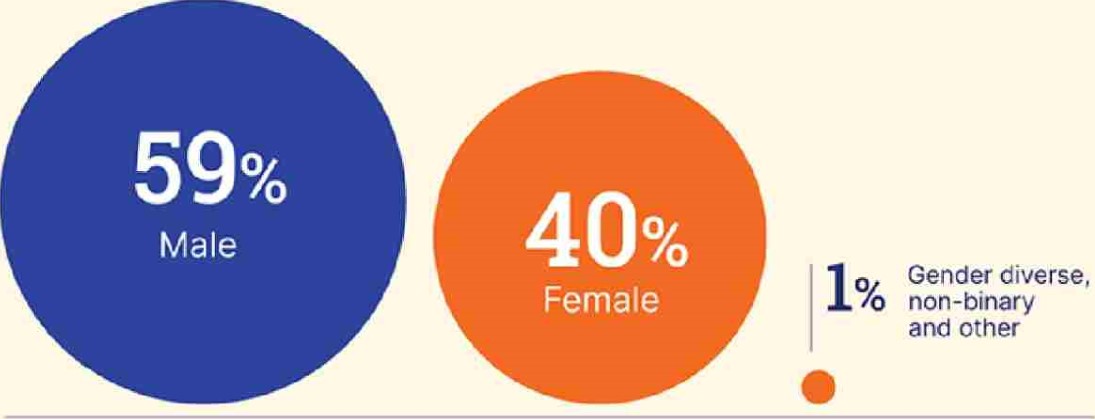
4. Key Insights

Key Insights

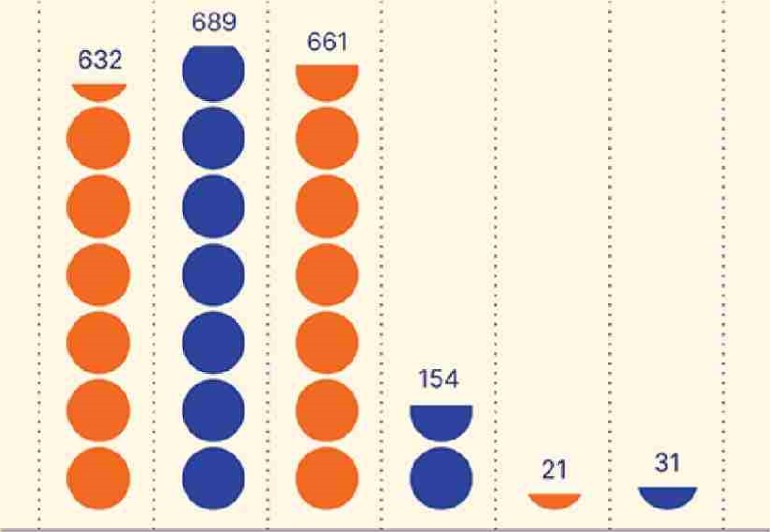


Insights 

|  |
| --- |
| GENDER |

Of the 2329 survivor accounts, 1378 or 59% were male, 932 or 40% were female, and 10 or 1% were gender diverse, non-binary or other.

The percentages refer to the proportion of people for whom we hold data, with 9 survivor accounts not including confirmation of gender.

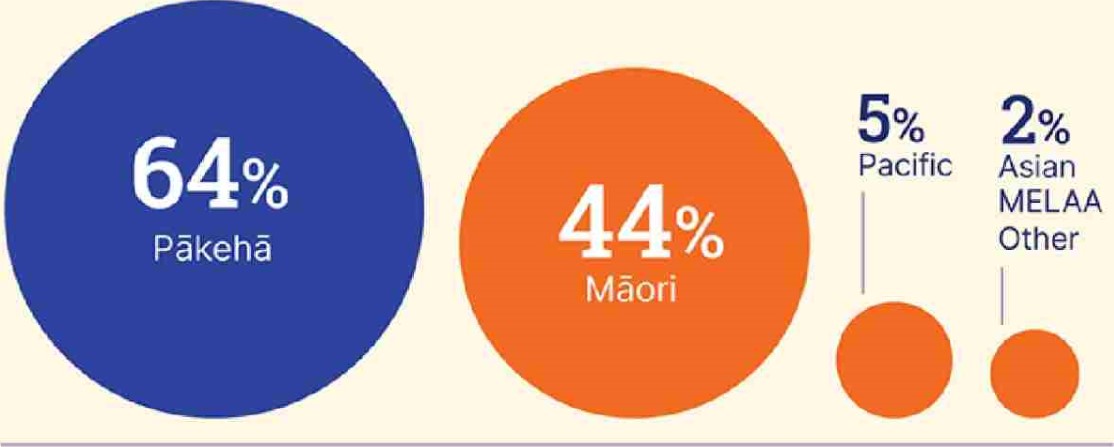
The survivor's age when first in care was identified for 2,188 (94%) of survivors. This includes survivors whose age, when first entering into caref was imputed from their setting.

0-4 5-9 10-14 15-19 20-24 25-64

AGE FIRST IN CARE O 100 survivors



Of the 2,329 survivor accounts, 2,233 provided ethnic identity.

The largest proportion of survivors identified as either Päkehä or Mäori. 1,483 or 64% of survivors identified as

Päkehä, while 44% identified as Mäori.

113 or 5% of survivors identified as Pasifika, 7 survivors identified as Asian and one survivor identified as Middle Eastern, Latin American or African (MEL AA). 53 survivors or 2% identified as Other ethnicities.

Unidentified: 4%

We note that survivors could identify with multiple ethnicities based on their family heritage, meaning the sum total of ethnicities is greater than the 2,233 survivor accounts.



For the purposes of our analysis we defined a gang member as a patched or prospective gang member. Of the 2329 survivor accounts, 229 survivors confirmed they were a member of a gang. Another 104 survivors stated they had a family or whänau member in a gang. Four survivors directly stated they were not in a gang. 1,992 survivors did not state or were not asked about gang membership and therefore we hold no information about them.



0

O

|  |  |  |
| --- | --- | --- |
| Of all survivor accounts, 29% of |  |  |
| survivors reported having been in |  | ooooooeee |
| prison at some point in their life | 29% | 0000 |
| (683 out of 2329 survivors). |  | eeeeeeeee |
|  | of survivors have | ooooooooe |
|  | been in prison at some point | ooooooooo |

EXPERIENCE OF INCARCERATION

|  |
| --- |
| MENTAL HEALTH |

83% of survivors reported living with mental distress at some stage during or following their state or faith-based care.of survivors

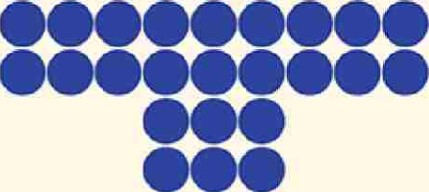
reported living with a mental health condition

|  |
| --- |
| CHRONIC HEALTH CONDITIONS |

000

Of all survivor accounts, our analysis that 33% of survivors 000

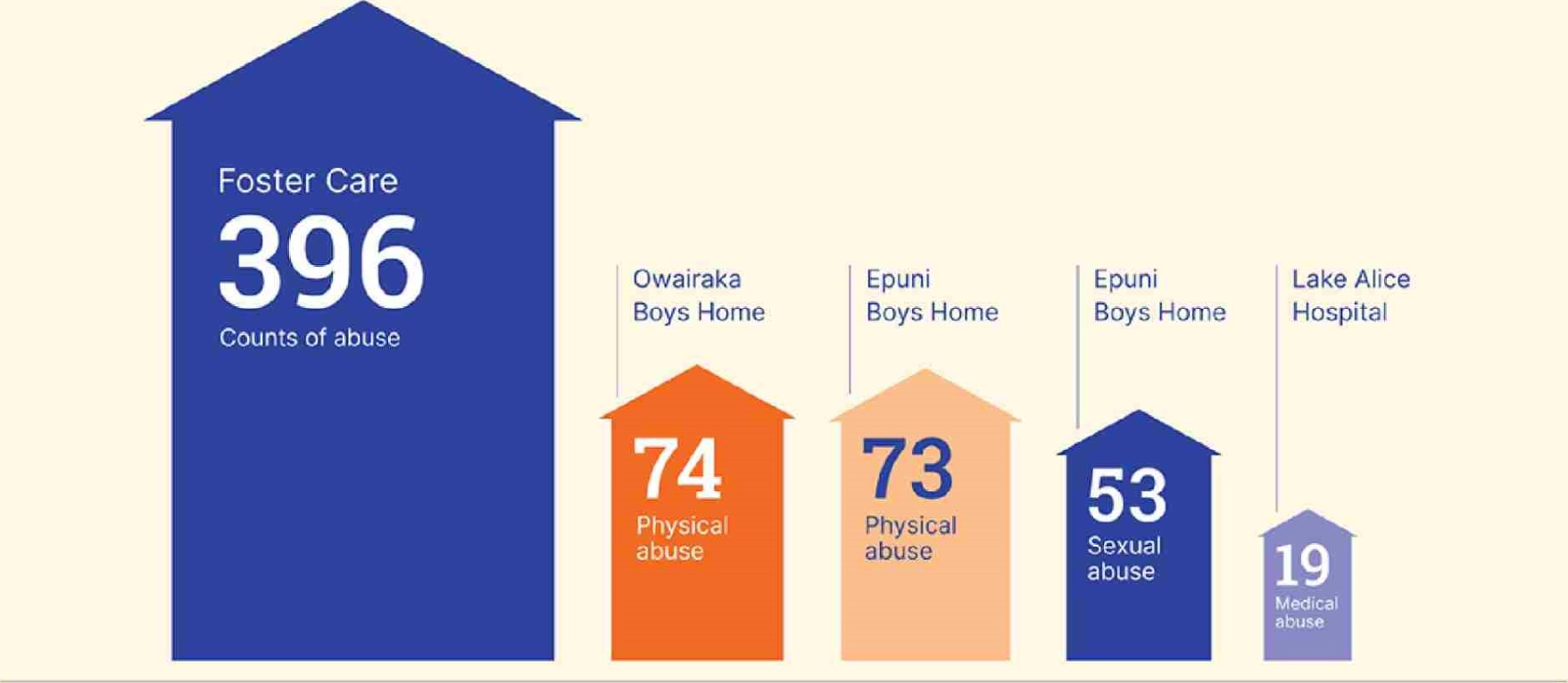
demonstrates reported living with a chronic health 33% 000 condition at some stage in their life ooooocooo

(769 out of 2329). of survivors reported living with a chronic health condition

## eee



|  |
| --- |
| ABUSE TYPE |

Total survivor counts by type of abuse.

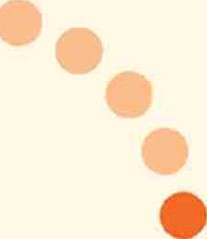


Non-contact



The graphic below shows the counts of abuse by type and institution. Foster care has the highest counts of abuse recorded across all abuse types except medical abuse, with 396 survivors reporting some form of abuse. Owairaka Boys Home and Epuni Boys home have the next highest counts of physical abuse at 74 and 73 respectively. 53 survivors reported sexual abuse at Epuni Boys home, while 19 survivors reported medical abuse at Lake Alice Hospital.

### Key

Of the 2,329 survivor accounts analysed, 1,398 identified their sexual orientation. 162 or 7% of survivors identified as LCBTQIA+. 1/236 or 53% C) 001 of survivors identified as not being LGBTQIA+J oo

while 41 or 2% of survivors preferred not to identify their sexual orientation.



Of the 1,018 Mäori survivors with accounts, 74% (755) reported an affiliation to at least one iwi. Here are the counts of iwi by the broadest Statistics New Zealand iwi categorisation (Iwi and iwi-related groups statistical classification V2.1.0).

Note that the total is larger than the number of Mäori who reported iwi affiliation because each individual may be associated with more than one iwi.

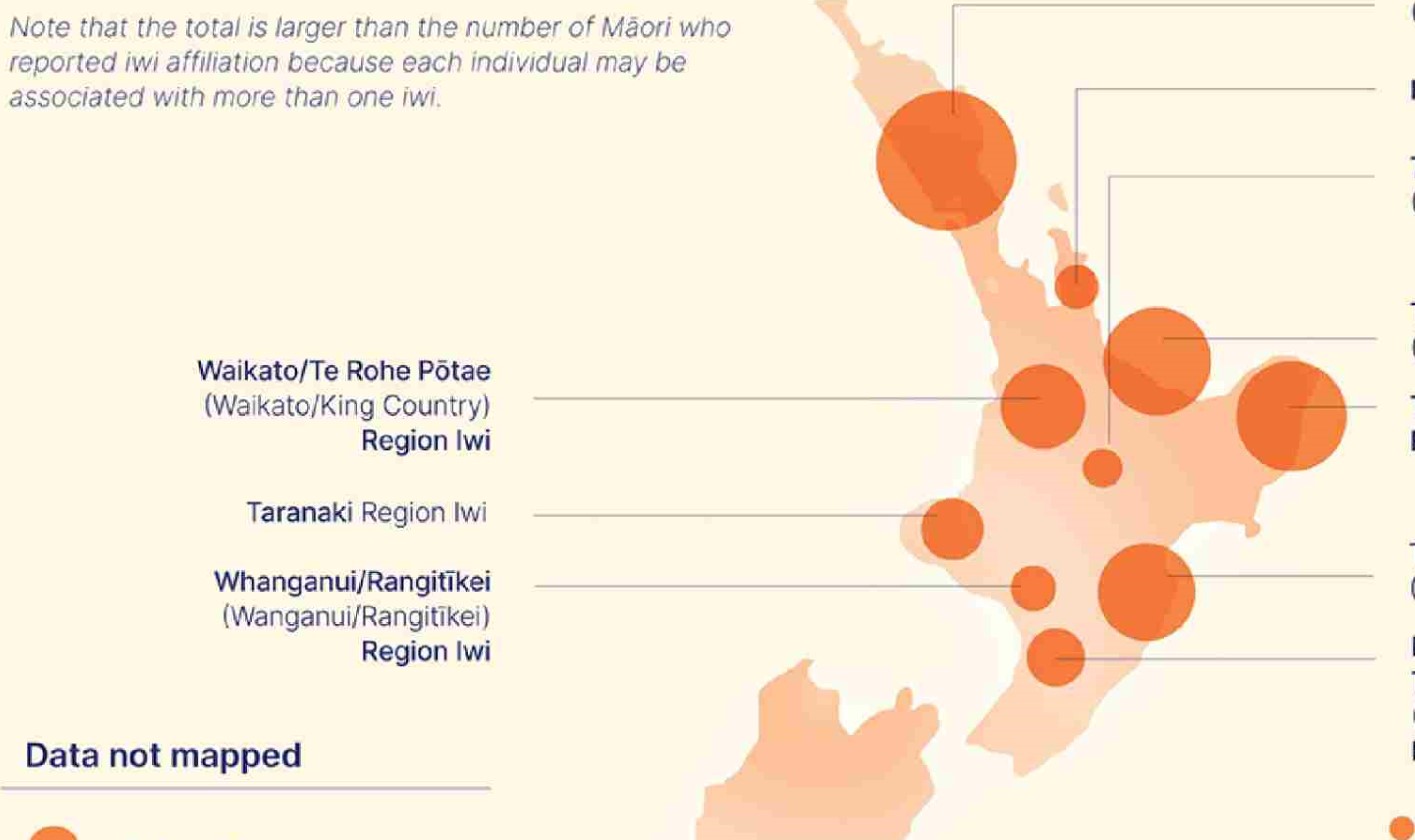
o Refused to answer

O Don't know

Iwi named, region not known

Confederations and Waka

o of survivors identified as LCBTQIA+

Te Tai Tokerau/Tämaki-makaurau (Northland/Auckland) Region Iwi

Hauraki (Coromandel) Region Iwi

Te Arawa/Taupö

(Rotorua/Taupö) Region Iwi

Tauranga Moana/Mätaatua (Bay of Plenty) Region Iwi

Te Tai Räwhiti (East Coast)

Region Iwi

Te Matau-a-Mäui/Wairarapa

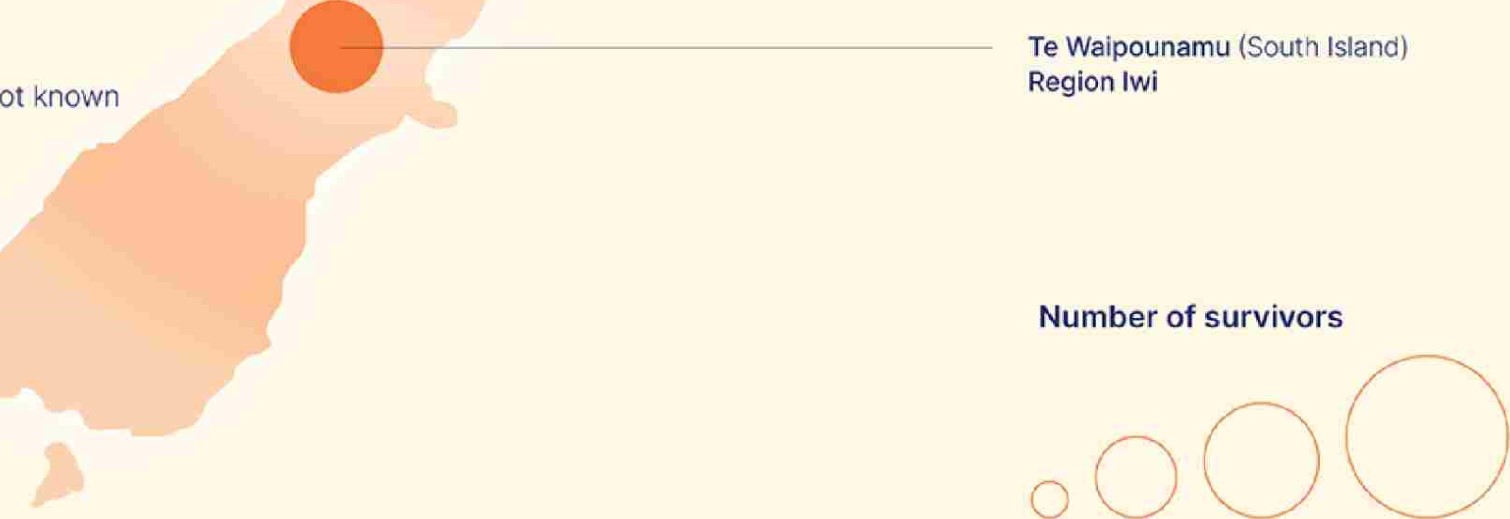
(Hawke's Bay/Wairarapa) Region iwi

Manawatü/Horowhenua/

Te Whanganui-a-Tara

(Manawatü!Horowhenua/Wellington) Region Iwi

Rékohu/Wharekauri

(Chatham Islands) Region Imi/lwi

10 50 100 200

### Key



UNIDENTIFIED 0.9% -—1 0.9%

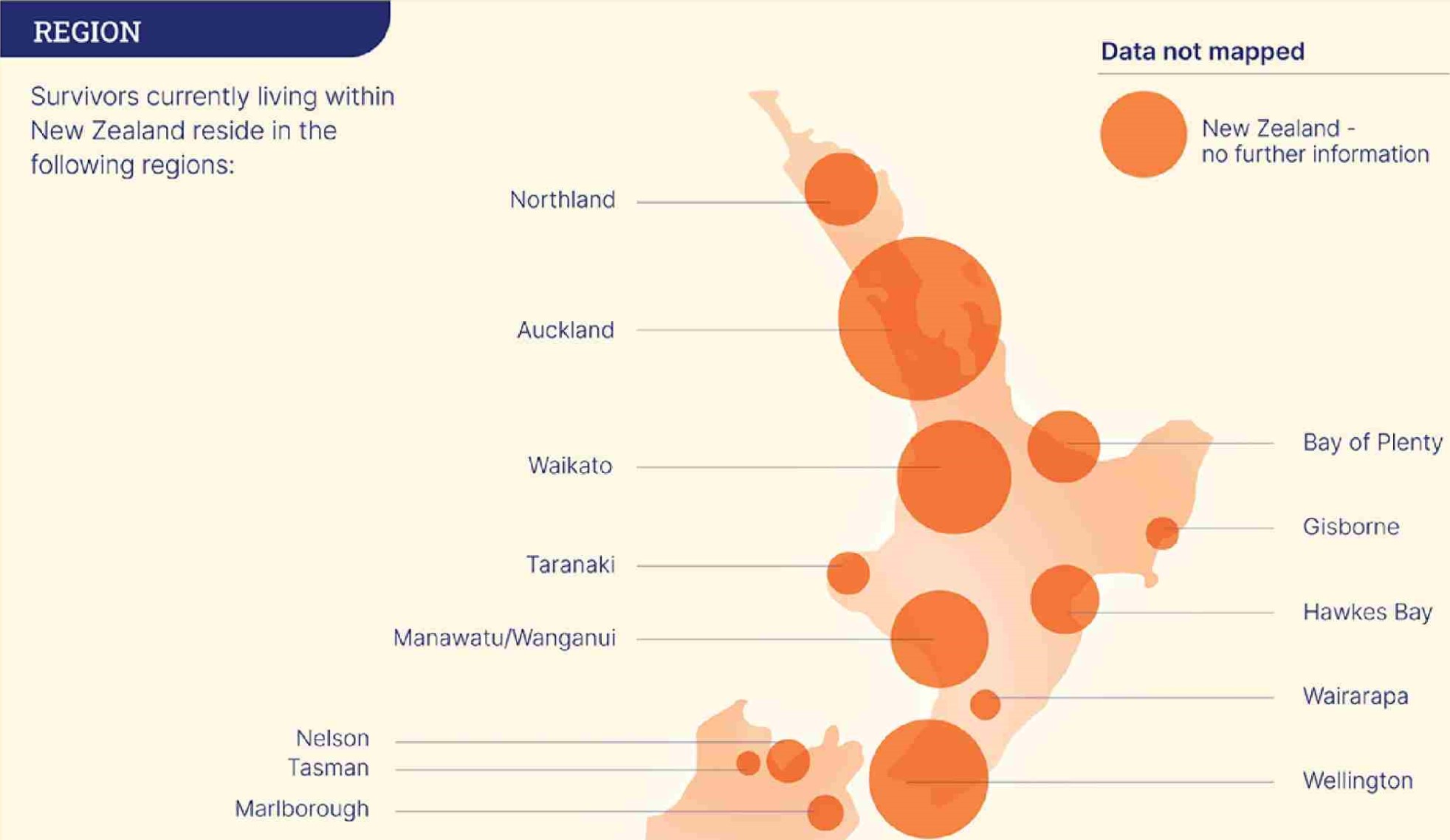
OTHER On this page we've detailed the reported region and country ofAUSTRALIA5% survivors based on each current survivor address or location where the witness statement is taken.

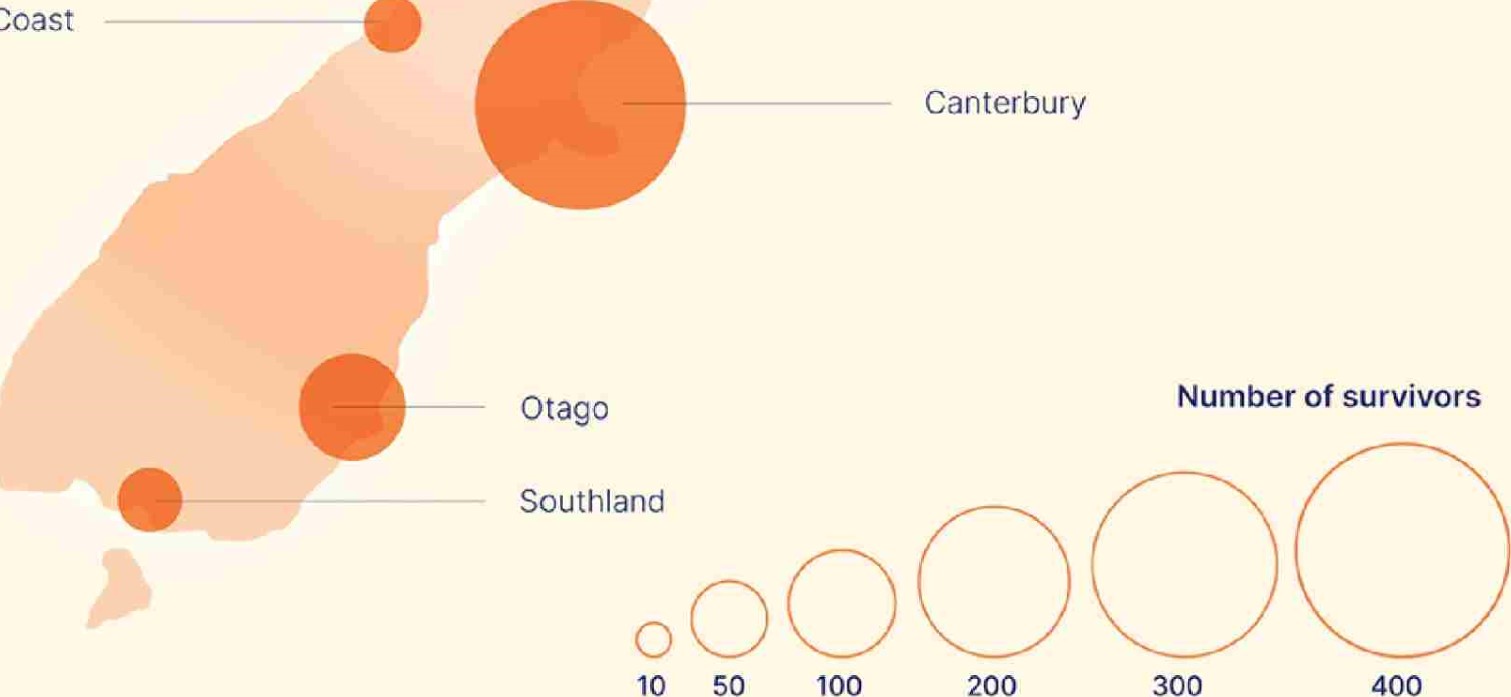


live

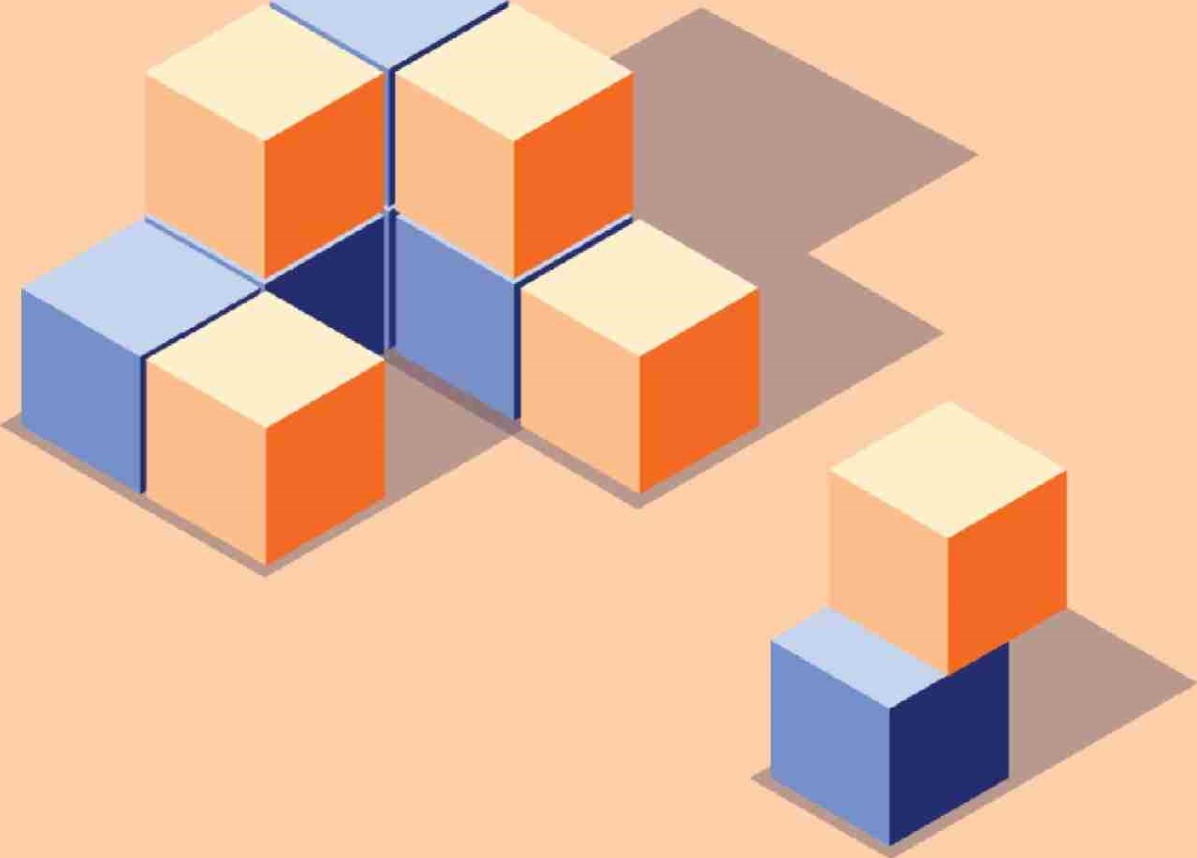
in

NZ



West Coast

# Demographic Analysis



5. Demographic Analysis

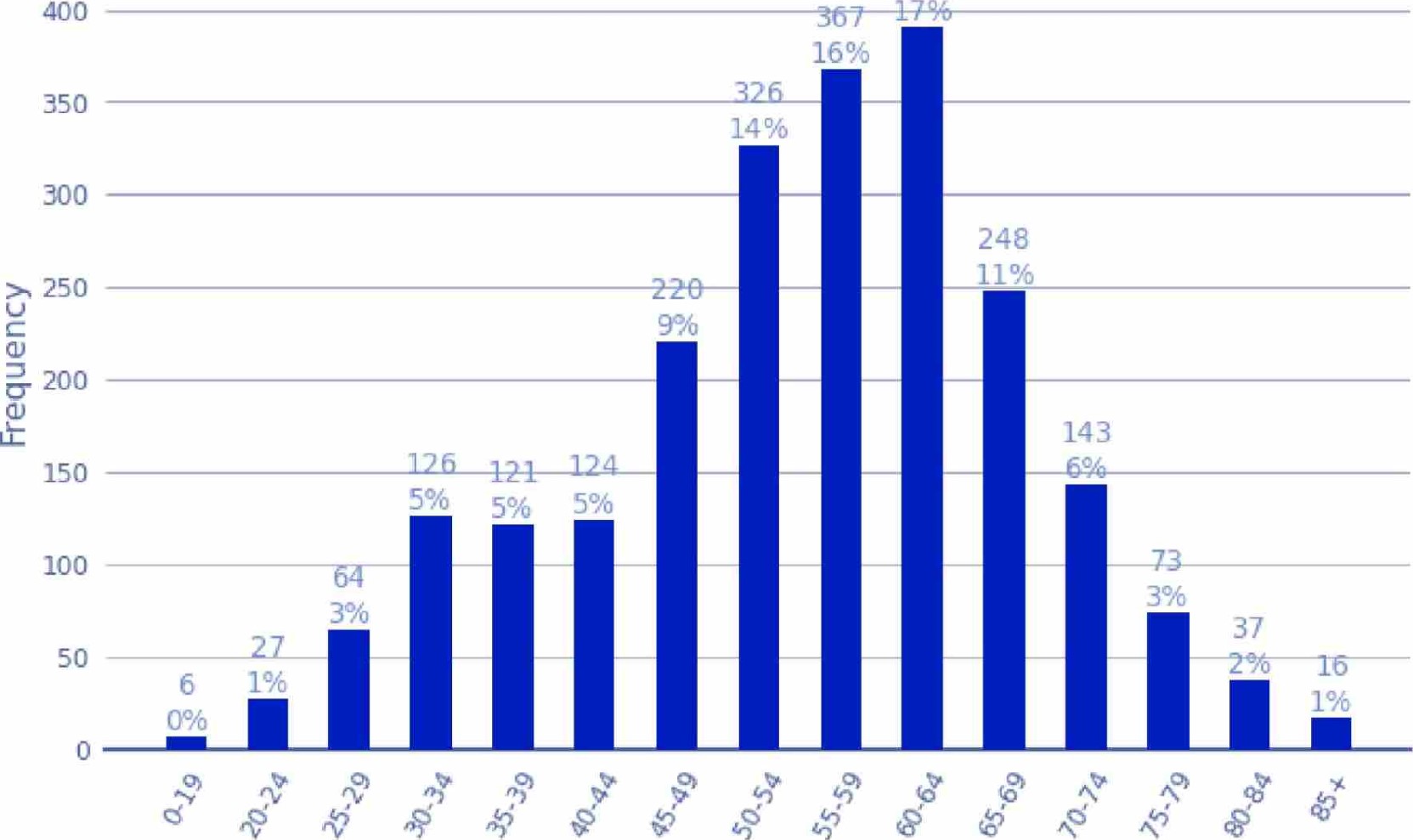
In this section of the report, DOT analysed the different demographic variables and cross-examined these to ga•n further insight of the vulnerability and propensity of certain demographic groupings.

## 5.1. Current Age of Survivors

When analysing the current age of survivors, close to half (47%) of survivors are aged between 50 and 64, with two-thirds (68%) being aged between 45 and 69. The largest grouping of survivors is aged between 60 and 64, but the age counts decline significantly after this bracket.

### Current Age of Survivors

390



#### Age

Figure 5: Number of Survivors by Age Group

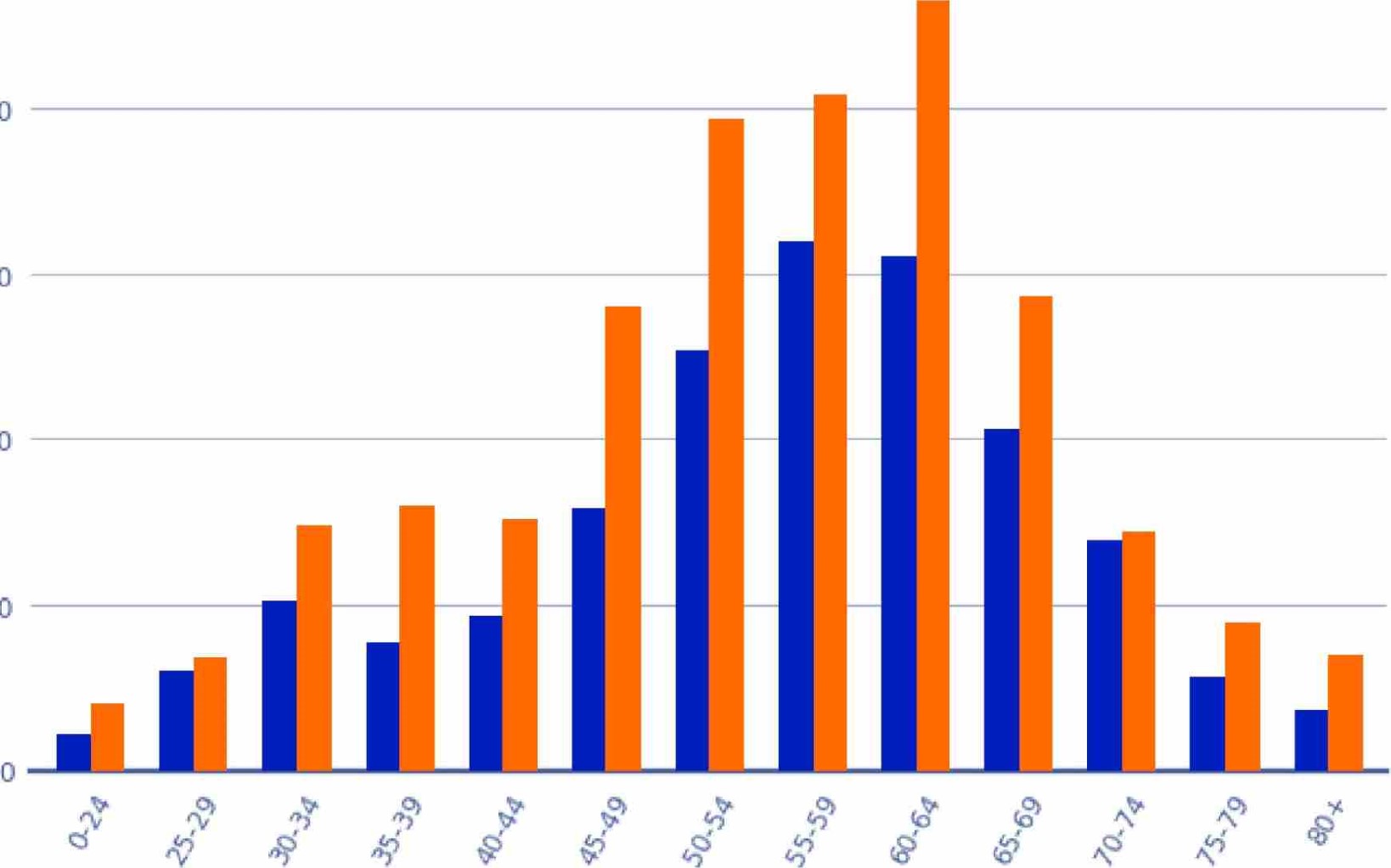
Please note the highest and lowest age brackets are wider than the rest, as they have been combined from other brackets for privacy purposes.

### 5.2. Current Age by Gender

For survivors' current age by gender, there are higher counts of males. This is reflective of 59% of total survivors being male, while 40% were female.

#### Current Age of Survivor by Gender

 Female Male

200

150

t: 100

50

Age

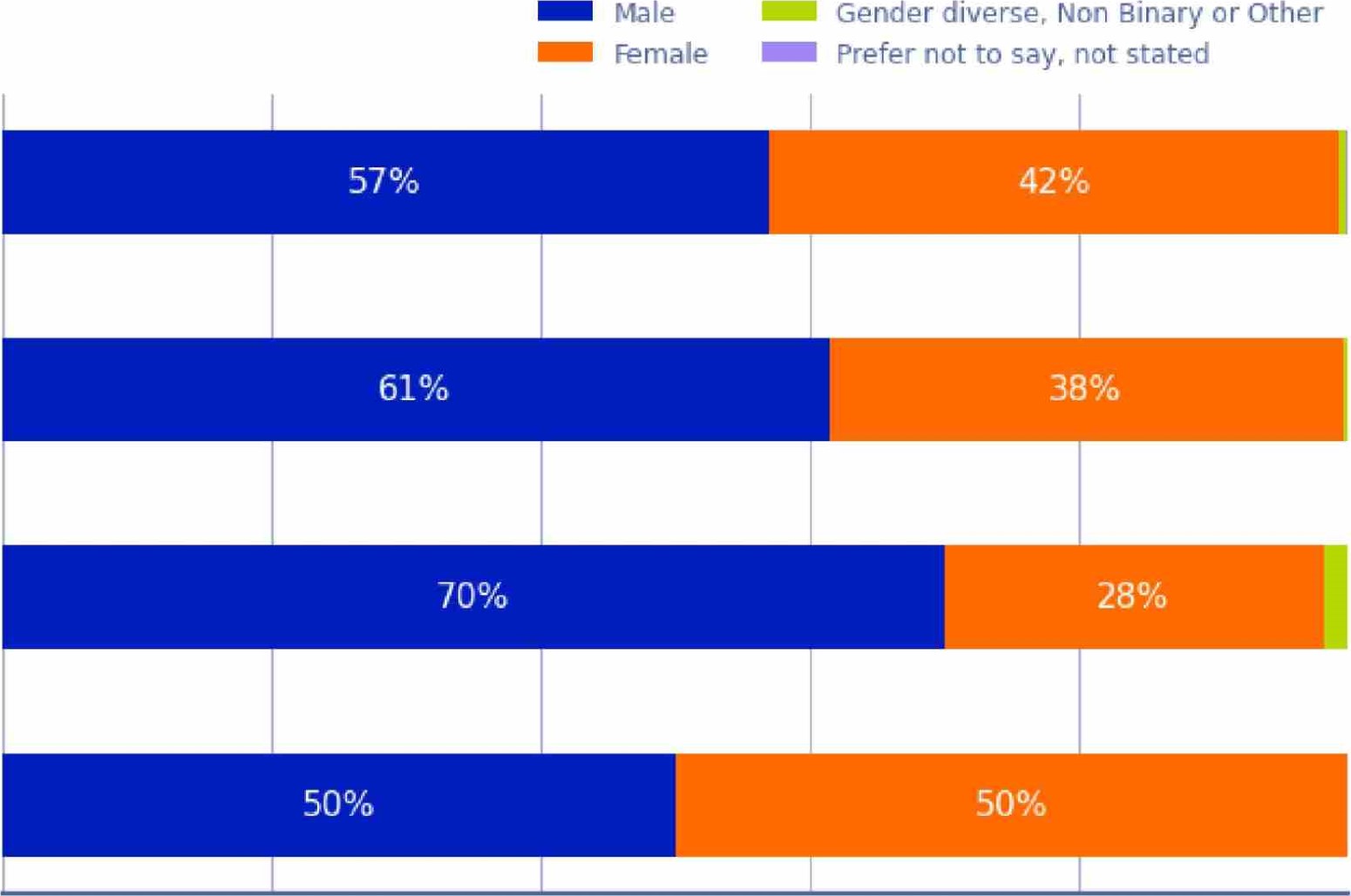
Figure 6: Current Age of Survivor by Gender

Please note that for this section only males and females are represented as the other gender types have been suppressed for confidentiality reasons.

## 5.3. Ethnicity by Gender

For each of the three main ethnicities, males reported higher levels of abuse. This reflects the higher count of male survivors in care.

### Proportion of Gender by Ethnicity



20%

60%

Proportion

of

Ethnicity

100%

Päkehä

Mäori

Pasifika

Other

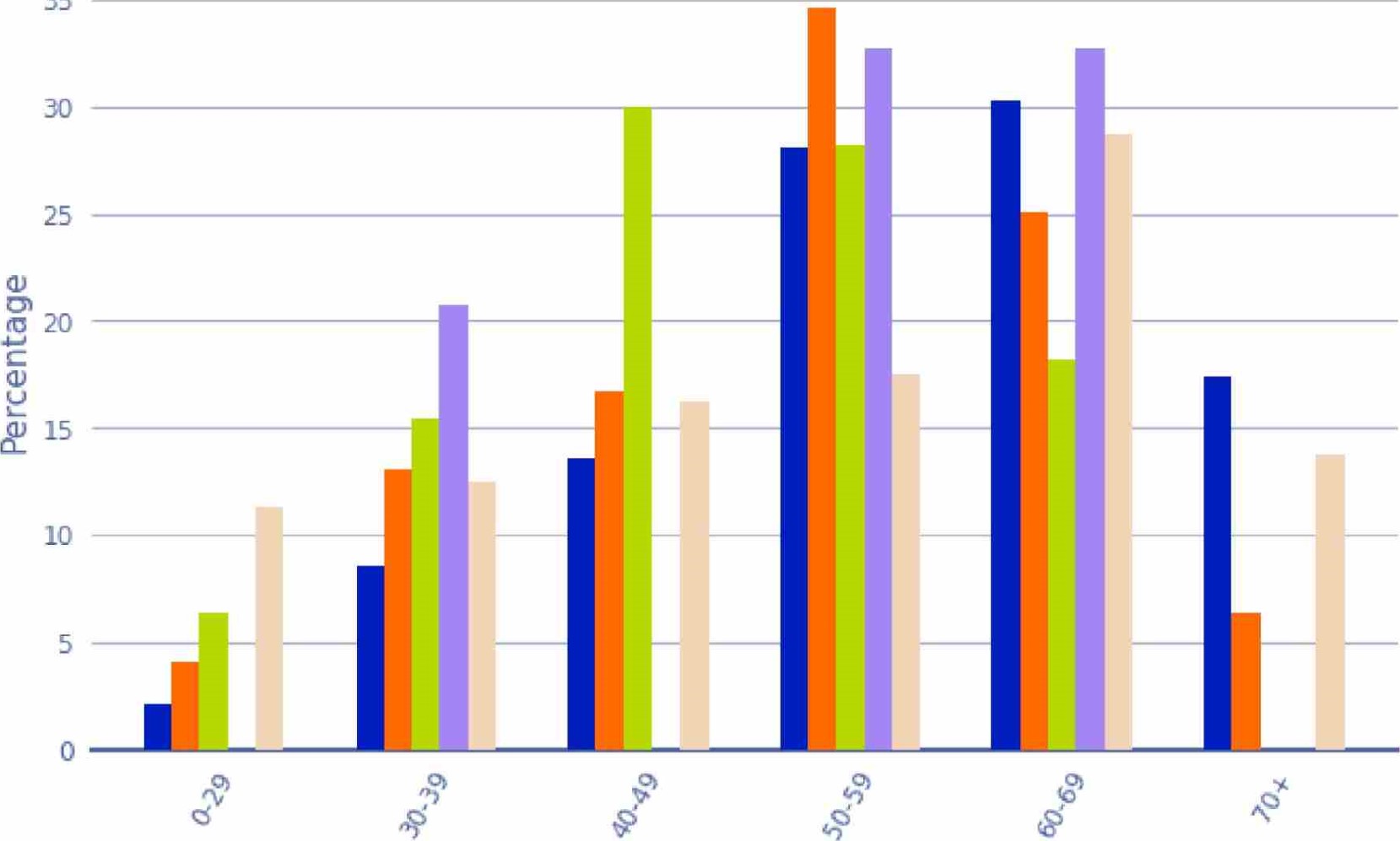
Figure 7: Proportion of Gender by Ethnicity

## 5.4. Current Age by Ethnicity

For current age by ethnicity, Mäori were most prevalent in the 50-59 age bracket, Pasifika were most prevalent in the 40-49 age bracket, and Päkehä most prevalent between 60-69.

### Current Age by Ethnicity

 Other Unidentified



35

#### Age Bracket

Figure 8: Current Age by Ethnicity

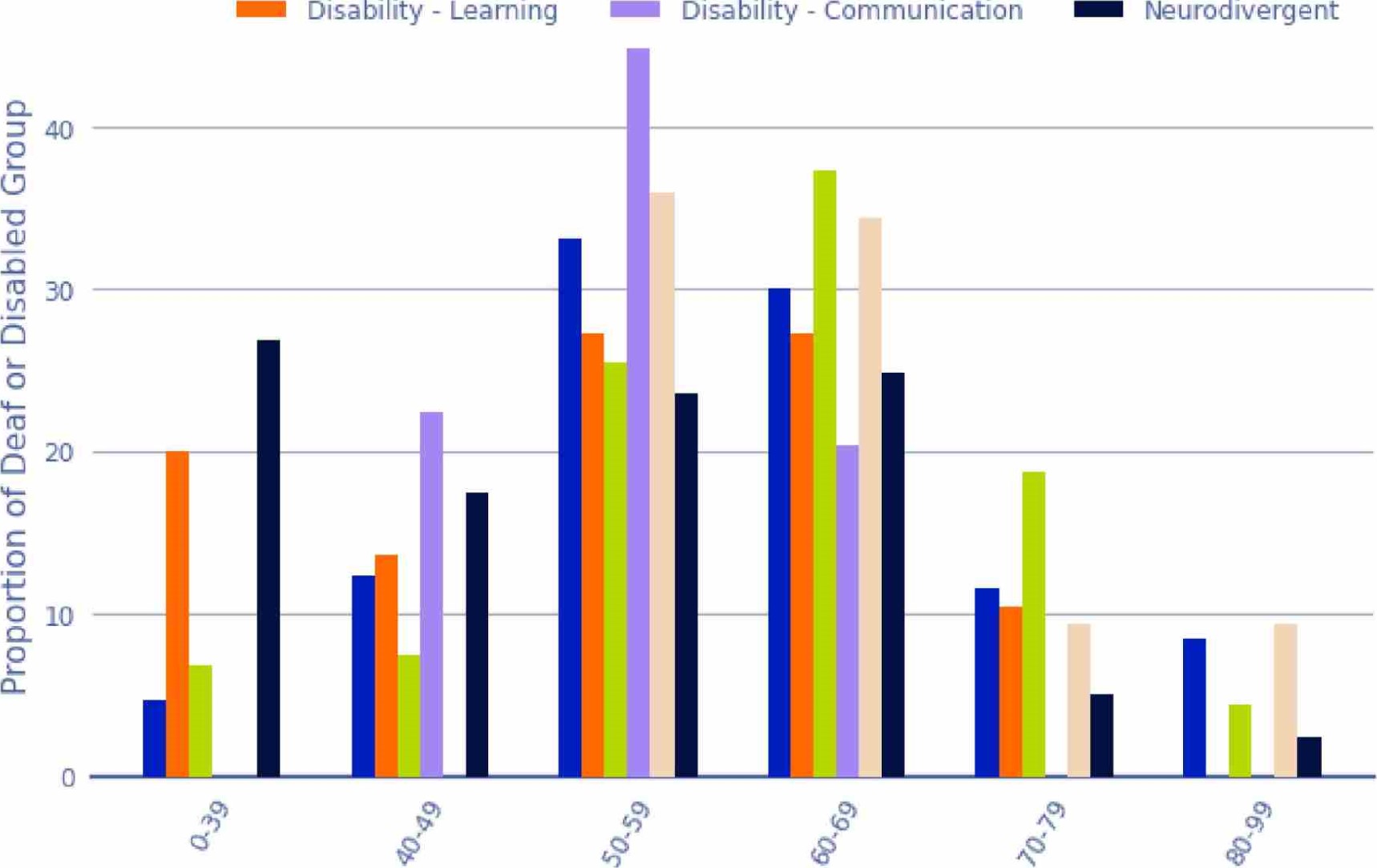
Please note that some counts for Pasifika and Other ethnicities were suppressed for confidentiality.

## 5.5. Deaf or Disability by Age

The age distribution of survivors who are Deaf or living with a disability closely reflects that for all other survivors.

### Deaf and Disabled by Age

 Disability - Mobility  Disability - Blind Disability - Communication Neurodivergent



Disability

-

Learning

#### Age

Figure 9: Deaf and Disabled by Age

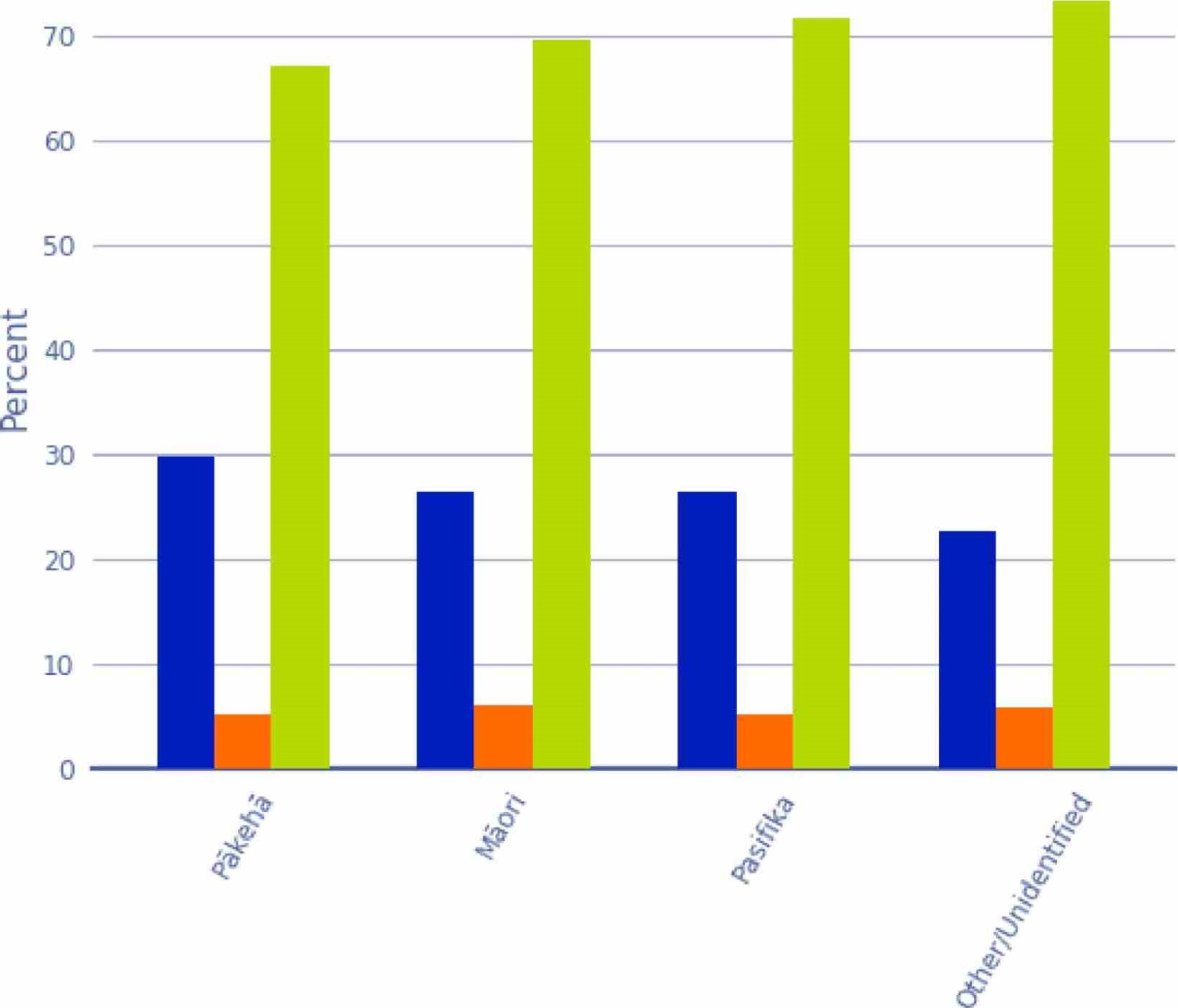
Note that some values for the Blind, learning, and communication disabilities were suppressed for confidentiality.

## 5.6. Deaf or Disability by Ethnicity

The graph below shows what percentage of each ethnicity are Deaf, living with a disability, or living without a disability. Across each ethnicity, the percentage of survivors who are Deaf, disabled or without a disability is consistent across the groupings.

Deaf and Disabled Survivors by Ethnicity

 Survivors with a disability Ckaf Survivors not Deaf or disabled



### Ethnicity

Figure 70: Deaf and Disabled Survivors by Ethnicity

The Other and Unidentified ethnicities were combined for confidentiality.

### 5.7. Experience of Incarceration

Of the 2,329 survivor accounts, 683 (29%) survivors were incarcerated at some point during their life.

#### 5.7.1. Incarceration by Gang Affiliation

Out of these individuals who had been incarcerated at some point during their life, 145 are or were gang members. 50 had or have gang affiliations through whänau, and 488 have no gang affiliations.

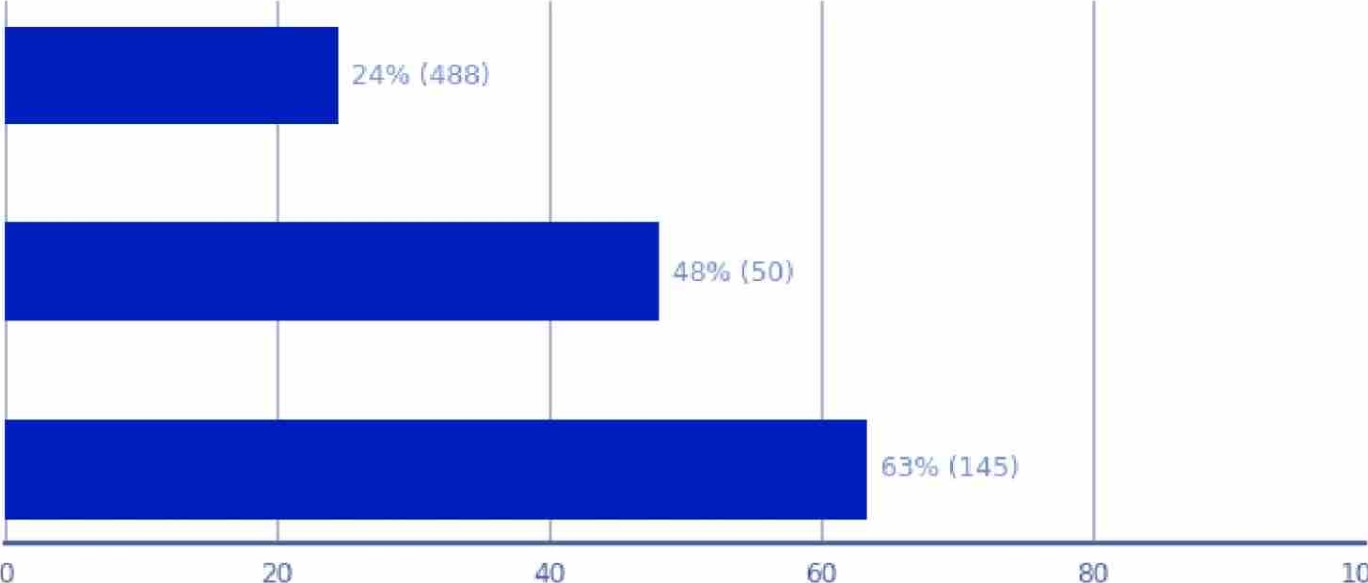
#### Survivor Incarceration by Gang Affiliation

No Association

o

E Through Whänau

Gang Member

100

Percentage

Figure 11: Survivor Incarceration by Gang Affiliation

##### 5.7.2. Ethnicity

The chart below details the percentage of survivors who have been incarcerated, by ethnicity. 42% of Mäori, 35% of Pasifika, and 24% of Päkehä survivors reported being incarcerated at some point during their life.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | 42% (425)  (40)  (23) |  |
| 0 | |
|  |  |

#### Survivor Incarceration by Ethnicity

Päkehä

Måori

Pasifika

Other

Unidentified

0 20 40 60

80 100

Percentage

Figure 12: Survivor Incarceration by Ethnicity

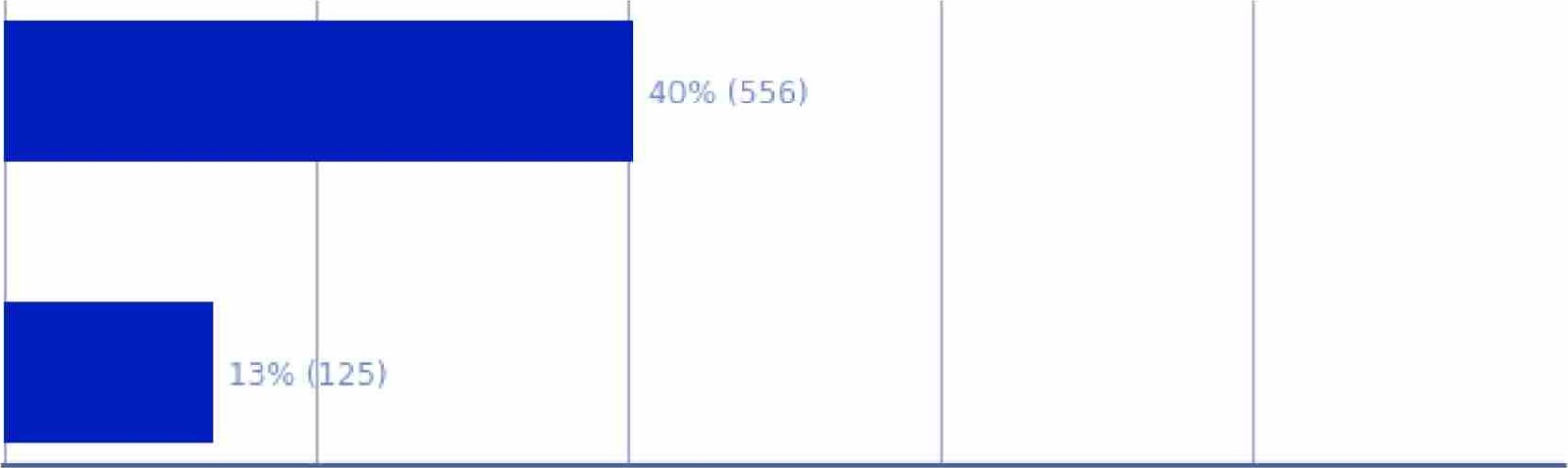
MELAA, Asian and Other ethnicities were combined into "Other" for confidentiality

5.7.3. Gender

The chart below details the proportion of the survivor population who have been incarcerated, by gender. Of these, 40% of male survivors and 13% of female survivors reported being incarcerated at some point during their life.

#### Survivor Incarceration by Gender

40% (556)



20

40

60

80

100

Female

##### Percentage

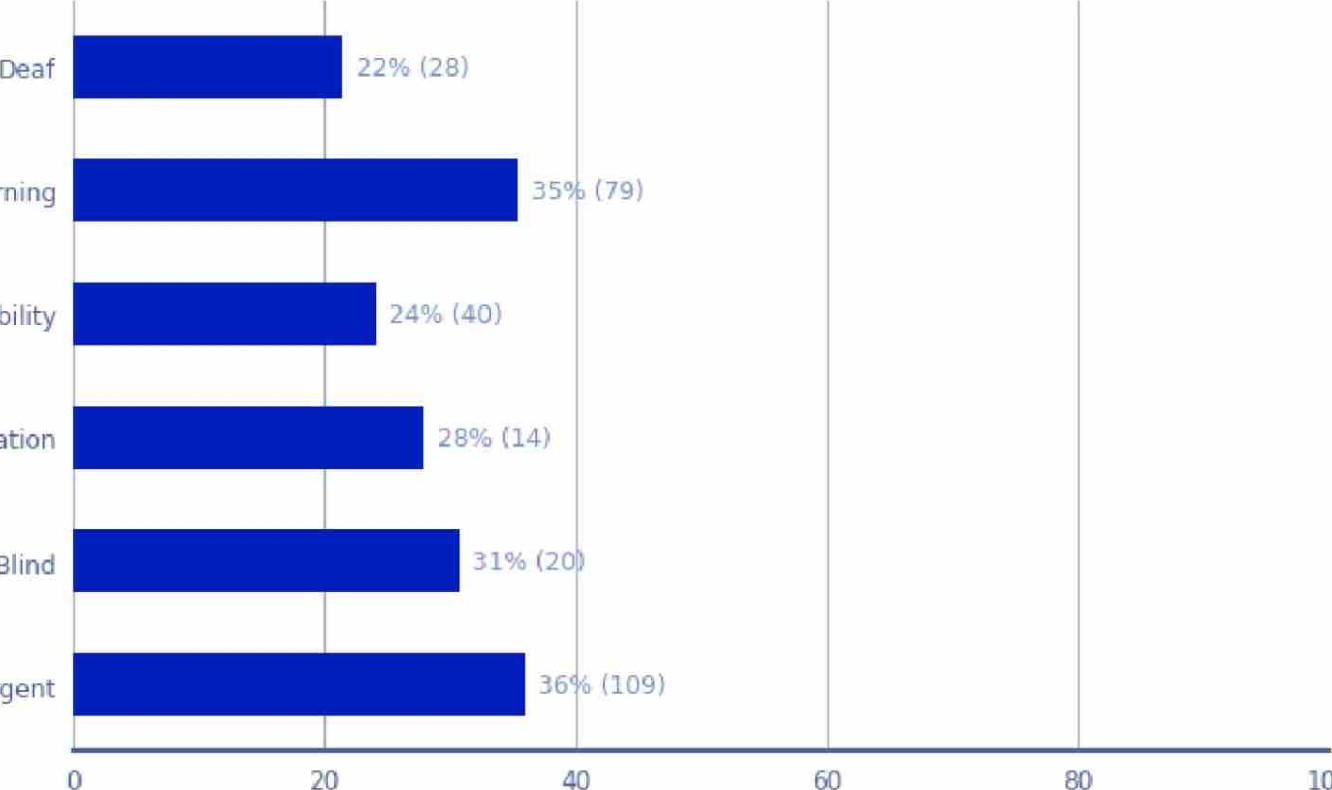
Figure 13: Survivor Incarceration by Gender

The "Prefer not to say" and "Gender diverse, Non Binary or Other" categories were suppressed for confidentiality.

##### 5.7.4. Deaf and Disability

The chart below shows the percentage of survivors who have been incarcerated at some point in their life and who also are Deaf or have a disability. Of the survivors ever incarcerated, 36% are Neurodivergent, 35% have a learning disability, and 31% are blind.

##### Survivor Incarceration by Deaf or Disability

Disability - Learning

Disability - Mobility

Disability - Communication

Disability - Blind

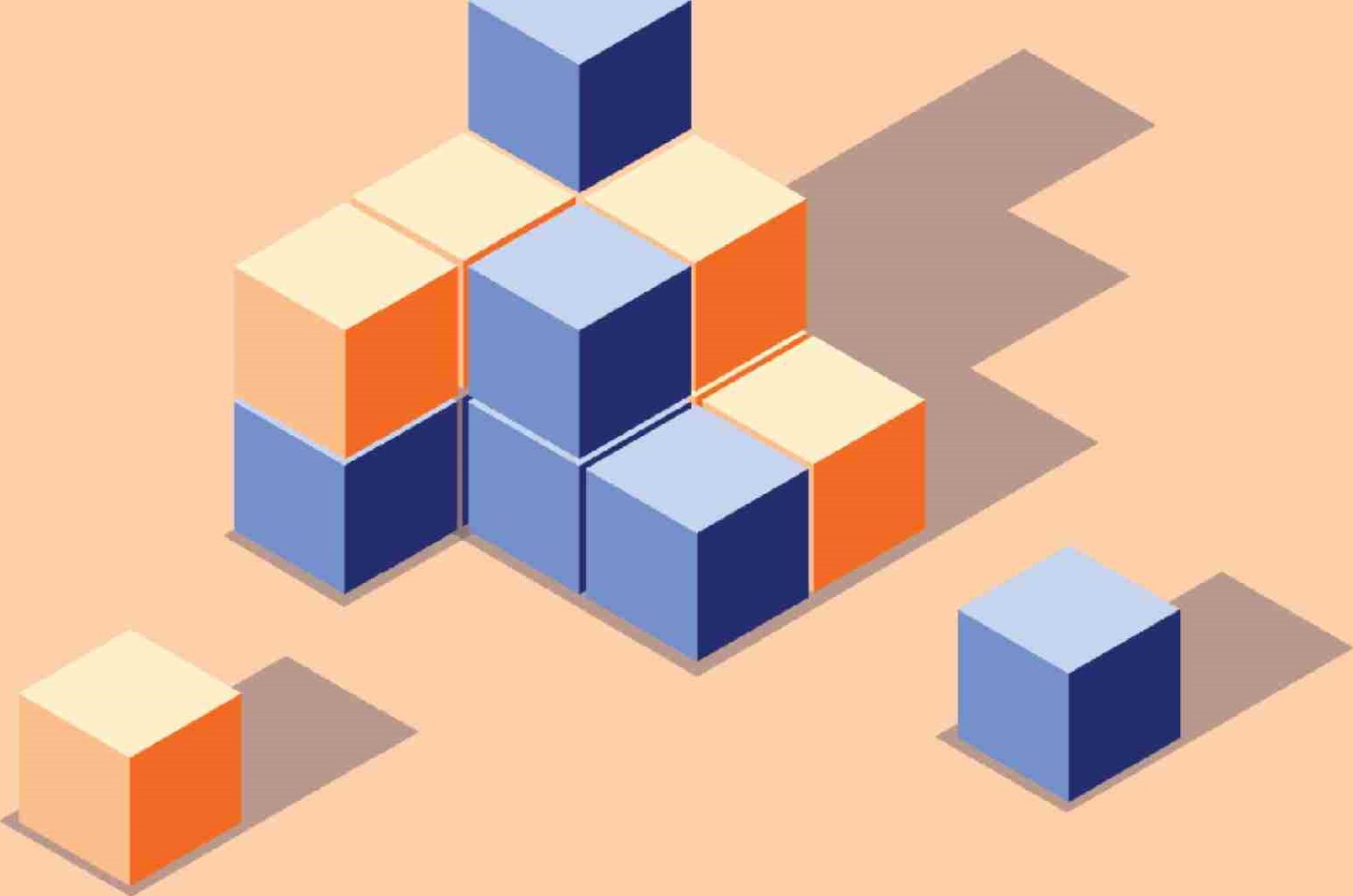
Neurodivergent

100

Percentage

Figure 14: Survivor Incarceration by Deaf or Disability

# Pathways into Care



6. Pathways Into Care

The survivor pathway into care is complex to determine. Often survivors do not know for certain why they were placed in state or faith-based care or prefer not to talk about it. Additionally, no data variable was consistently collected on survivors' pathways into care.

However, from the data we have identified pathways into care for 16% of survivors using natural language processing.

## 6.1. Findings for Pathways into Care

Almost half (46%) of survivors, for whom the pathway into care information was identified, were placed into care by state requirement. This can stem from a family being reported to (and investigated by) Child, Youth and Family Services (CYFS) due to suspected neglect or an unsafe environment for the children, or other reasons such as a misbehaviour" of a child or adolescent - for example, truancy from school, stealing or running away from home.

Close to a quarter (24%) of survivors with a pathway reported being placed into care due to their "parents' voluntary placement" with state or faith-based organisations.

Other typical pathways include:

* Enrolment into faith-based boarding or day schools (9%).
* Placement into a psychiatric care setting (7%) due to suspected mental health problems or hard-to-manage behaviour.
* An abrupt change in family circumstances (4%), leading to a forced or voluntary placement of a child or children with CYFS or another state or faith-based organisation. For example, this may have been due to the death of a parent/ illness or divorce.

## 6.2. Categories of Pathways Into Care

Survivors' pathways into care were classified using the following categories:

State requirement - Typically through the Court or CYFS for either:

 Care and protection issues: all types and forms of abuse and neglect at home

* Youth justice-related issues: usually lower-level offending like stealing food, car theft  Issues related to 'delinquency': not under proper/adequate control, acting out, playing up, delinquent, indigent, getting into trouble, naughty, mischief, promiscuous, sexual promiscuity, running away from home or truancy/wagging from school.
* Reasons unknown to the survivor.

Parents' voluntary placement - Voluntary placement of a newborn or a child into CYFS care usually due to:

* Unwanted child/result of violence
* Not managing financially (poverty) or a lack of adequate support
* Not able to manage child behaviour.

Tragedy - Parental death/divorce/mental illness, abrupt change in circumstances leading to either no one left to care for a child or the remaining parent(s) not being able to take care of the child.

Faith-based school - Voluntary enrolment of a child into a faith-based school (boarding or day school).

Psychiatric - Admission to psychiatric settings due to:

* Troubled behaviour
* Homosexuality
* Depression
* Other conditions.

Disability - Admission to disability settings, including special and residential schools.

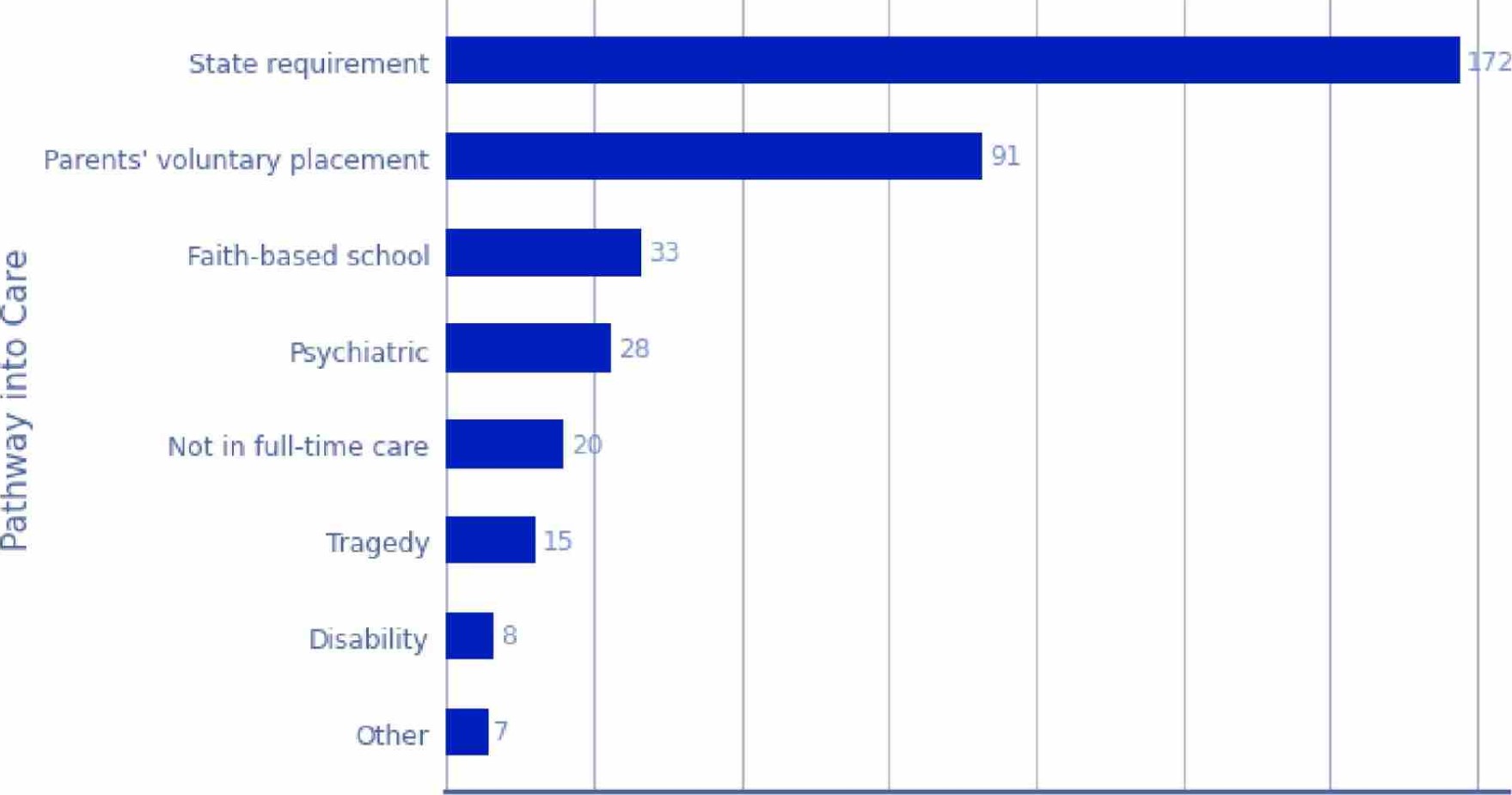
Not in full time care - Survivor does not have a pathway into care, because they were never in a full-time state or faith-based care, but they reported abuse at:

* Church - by influential figures in church
* State (day) school
* State health services

Faith community - Born into or brought as a child into one of the Faith communities

Unwed pregnancy - Pregnancy (pathway into unmarried mothers' homes, particularly for single, young women and girls).

Frequency of Pathways into Care



25

50

75

100

125

150

175

Frequency

Figure 15: Frequency of Pathways into Care

The 'Other' category includes both survivors who got into care through Unwed Pregnancy and those from Faith Communities; the categories are merged for confidentiality.

## 6.3. Pathway into Care by Decade of Abuse

The below analysis looks at typical pathways into care by the decade in which the reported abuse in care occurred. Imputed values for the decade of abuse are used, where the actual information is missing (see appendix for details about the imputation methodology).

In the figure below, any values less than 6 are suppressed, "unknown" category is excluded.

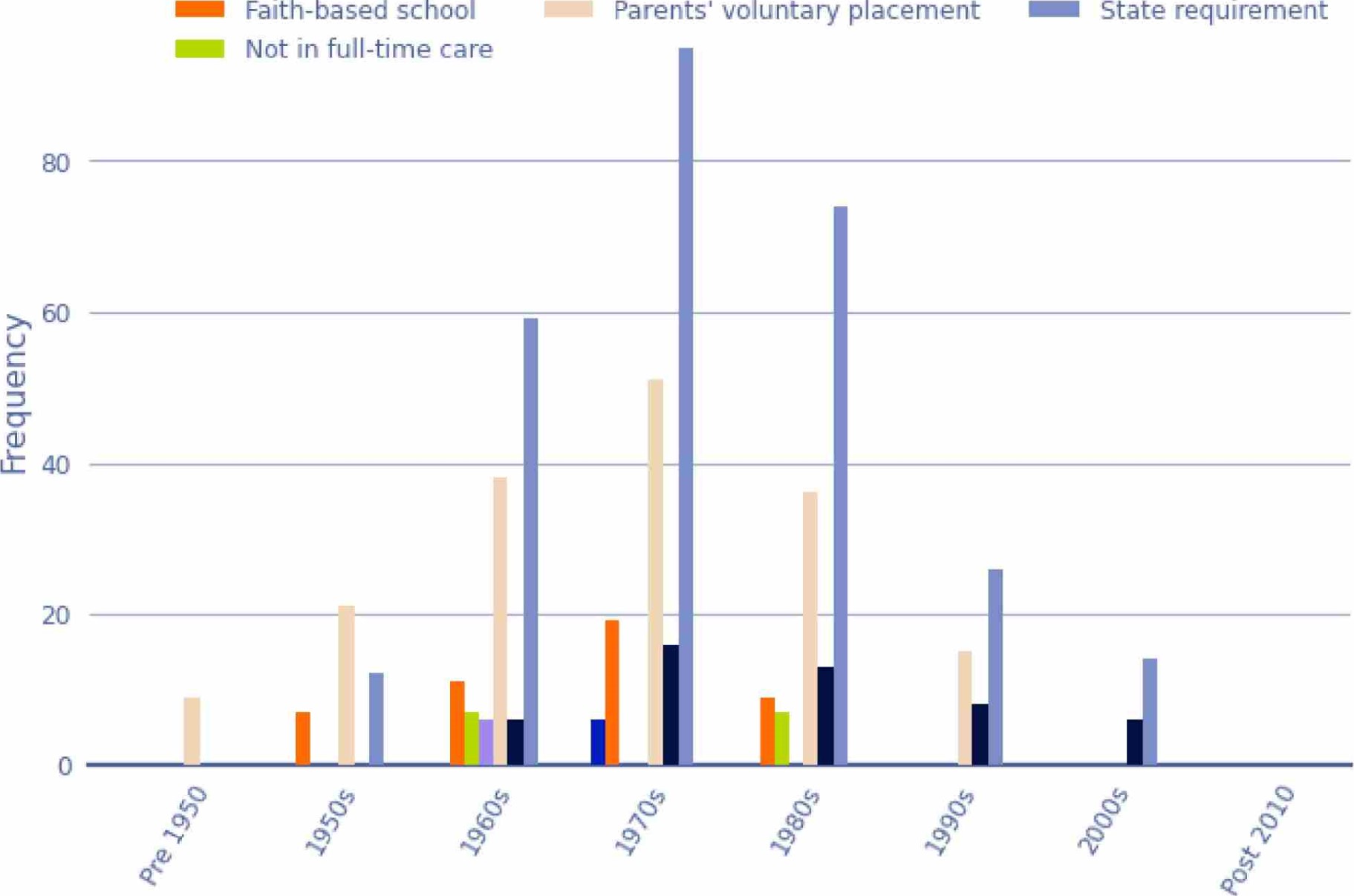
## Pathway into Care by Decade of Abuse

Psychiatric



Disability

Tragedy

State requirement

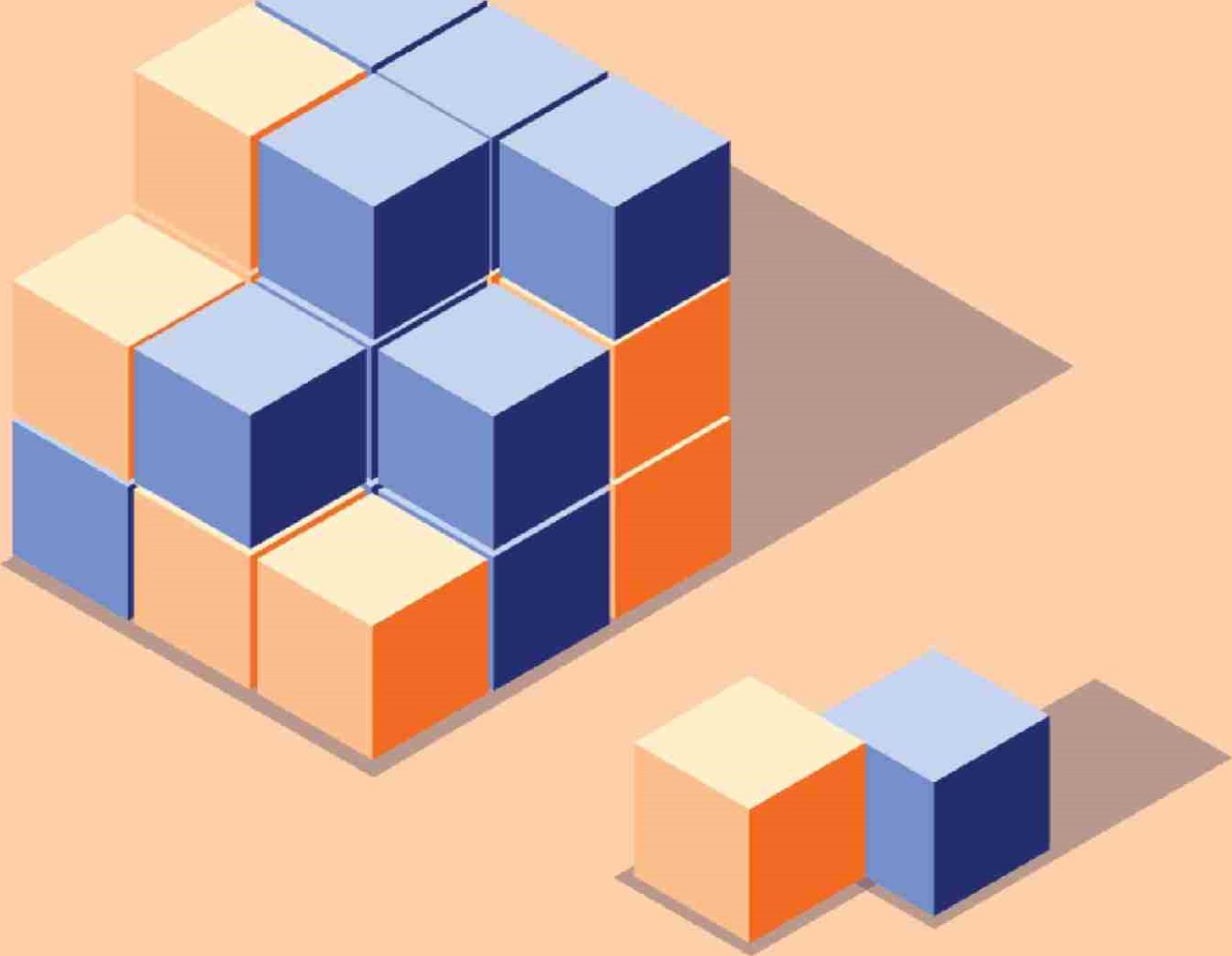
### Decade of Abuse

Figure 76: The Distribution of Pathways into Care by Decade of Abuse

Based on the survivor reports, 'parent's voluntary placement' with social services was a more typical pathway into care pre-1950s and in the 1950s. Starting from the 1960s, involvement of CYFS to initiate the transition to state or faith-based care increased.



# Abuse Types



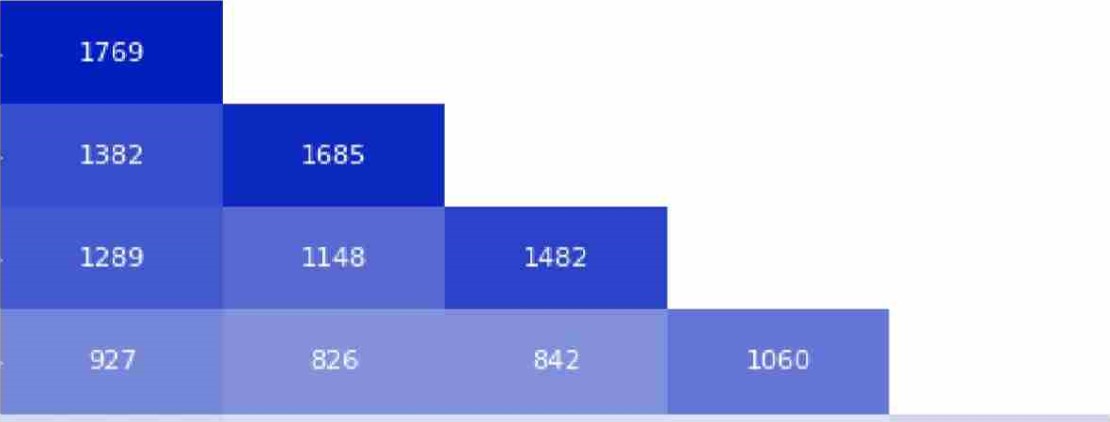
7. Abuse Types

Of the 2,329 survivor accounts, 2,168 reported some form of abuse. The abuse type and its prevalence varied depending on the decade, setting, institution type and the survivor. It was common for survivors to be subjected to multiple types of abuse, including but not limited to emotional, physical and sexual abuse, as well as abuse by neglect. In the following sections we have detailed our analysis of abuse types by decade of abuse, gender, age, ethnicity, Deaf and disability.

## 7.1. Relationship between Abuse Types

Below we've depicted the relationship between abuse types, which shows that many survivors reported multiple forms of abuse. The table observes that 1,382 survivors reported both physical and sexual abuse, while 1,289 survivors reported emotional and physical abuse and 1,148 survivors reported emotional and sexual abuse.

Relationship between Survivor Reported Abuse Types

Physical

Sexual

Emotional

Neglect

Medical 250 212 209 192 279

Non-Contact 164 179 150 119 30 201

Sexual

Solitary 196 166 171 189 60 20 213

Physical Sexual Emotional Neglect Medical Non-Contact Solitary

Sexual



200 400 600 800 1000 1200 1400 1600

Survivor Abuse Counts

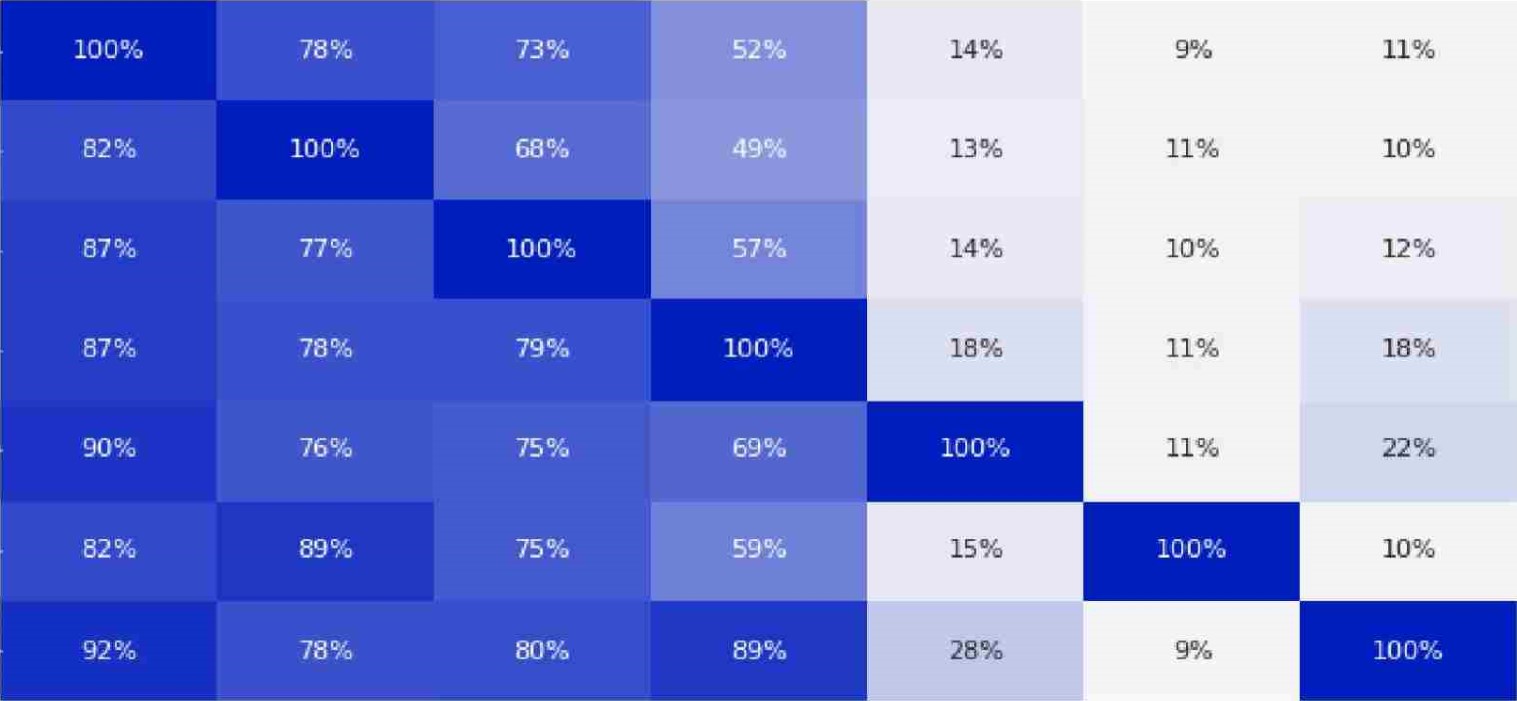
Figure 17: Relationship between Survivor Reported Abuse Types

## 7.2. Relationship between Abuse Types

The chart below details the relationship between abuse types to highlight the multiple abuse types survivors were subjected to. It is similar to the chart above, but shows proportions. For example, 82% of survivors who reported sexual abuse also reported physical abuse. 78% of survivors who reported physical abuse also reported sexual abuse.

The chart also shows that of those who reported solitary abuse, 92% reported physical abuse and 89% reported neglect,

Relationship between Survivor Reported Abuse Types

Physical

Sexual

Emotional

Neglect

Medical

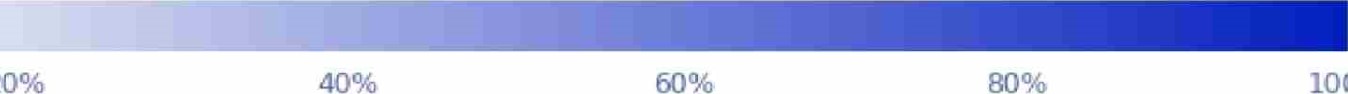
Non-Contact

Sexual

Solitary

Physical Sexual Emotional Neglect Medical Non-Contact Solitary

Sexual



20%

100%

Proportion

Figure 18: Relationship between Survivor Reported Abuse Types

## 7.3. Decade of Abuse by Abuse Types

The below analysis looks into the decade in which reported abuse occurred, by the abuse type. The scope of this analysis only includes survivors with accounts from 1950 - 1999, however as the data also has records in decades outside of this timeframe, we have chosen to include these in the analysis and have grouped the out of scope accounts as "Pre 1950" and "Post 2010".

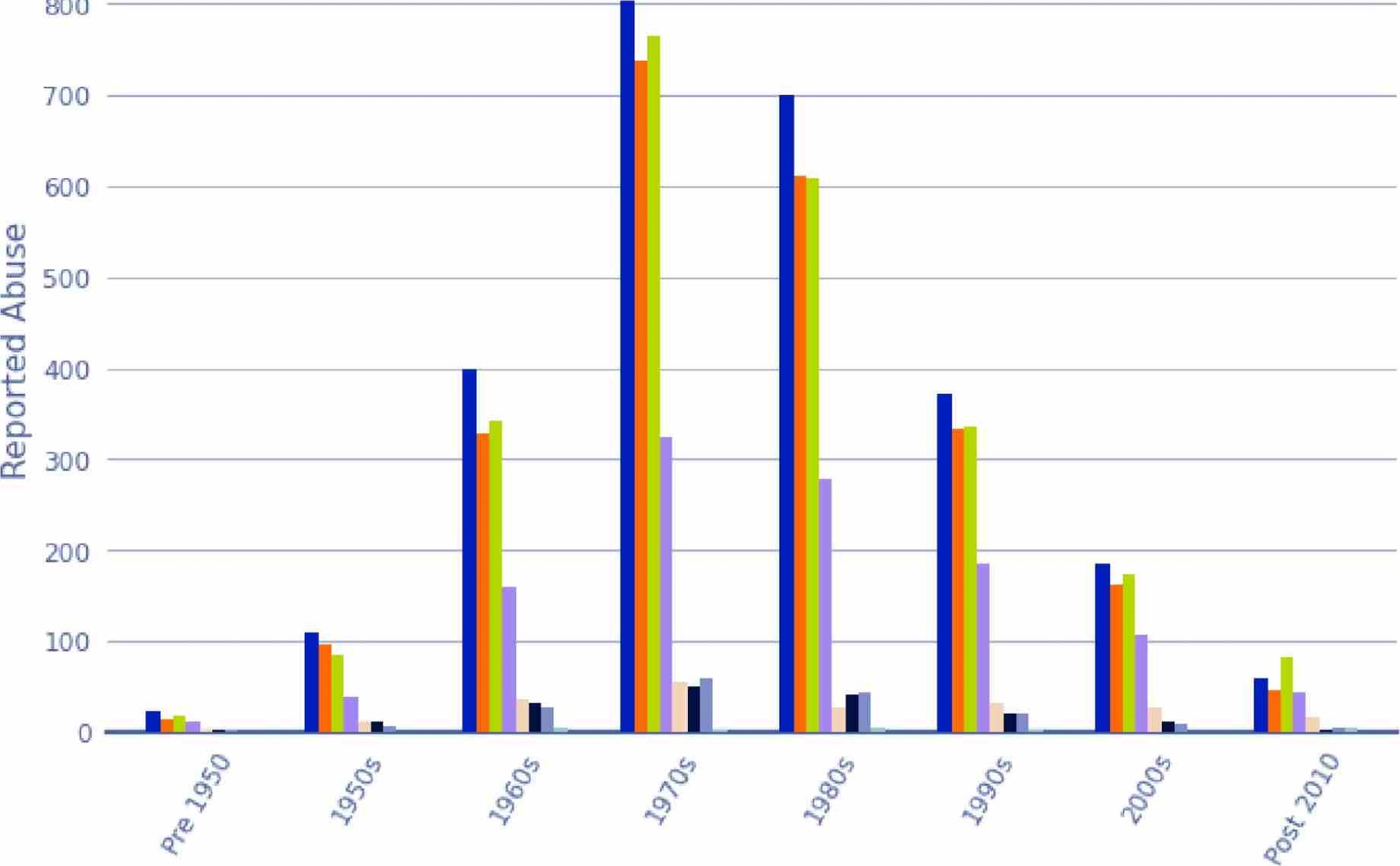
The decade of abuse was absent for 20% of the abuse data. To ensure the fullness of reporting, the decade of abuse was imputed based on the survivor age while in care (if available) and otherwise based on the setting. Further detail on this methodology can be found in the appendix.

### Frequency of Abuse Reported by Decade and Type of Abuse

 Physical  Neglect Medical

 Sexual  Forced or Child Labour Solitary Confinement

 Emotional/Psychological  Sexual (non-contact)



800

#### Decade of Abuse

Figure 79: Frequency of Abuse Reported by Decade and Type of Abuse

The three most prevalent abuse types reported by survivors for the period between 1950 and 2000 were physical abuse, sexual abuse and emotional abuse. These types of offences accounted for 5,439 or 82% of the instances of abuse over the total period.

Based on the survivor reports, the highest decade of abuse was the 1970s, with 11392 survivor reports of abuse in this decade. This compares with 713 instances in the 1960s and 996 in the 1980s.

## 7.4. Distribution of Abuse Types by Decade of Abuse

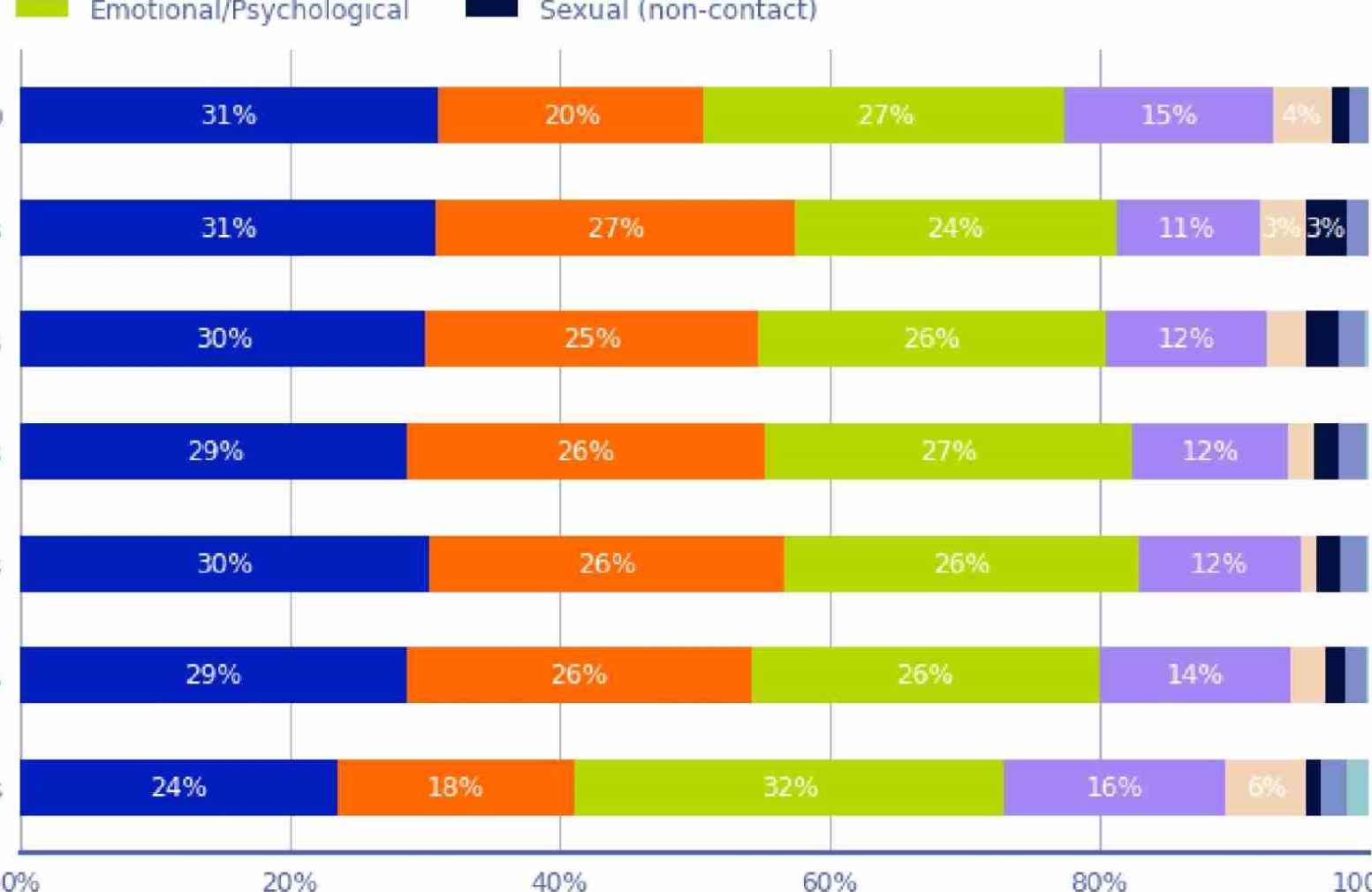
When we analysed the proportion of abuse type by decade, it was found that the abuse types remained largely consistent across all observed decades. Across all decades, physical abuse is the most prevalent, except for in post 2000 when accounts of emotional and

psychological abuse climbed to 32% of all reported abuse.

### Distribution of Abuse Type by Decade

|  |  |  |
| --- | --- | --- |
| Physical | Neglect | Medical |
| Sexual | Forced or Child Labour | Solitary Confinement |

Sexual (non-contact)



80%

Pre 1950

1950s

1960s

1970s

1980s

1990s

Post 2000s

100%

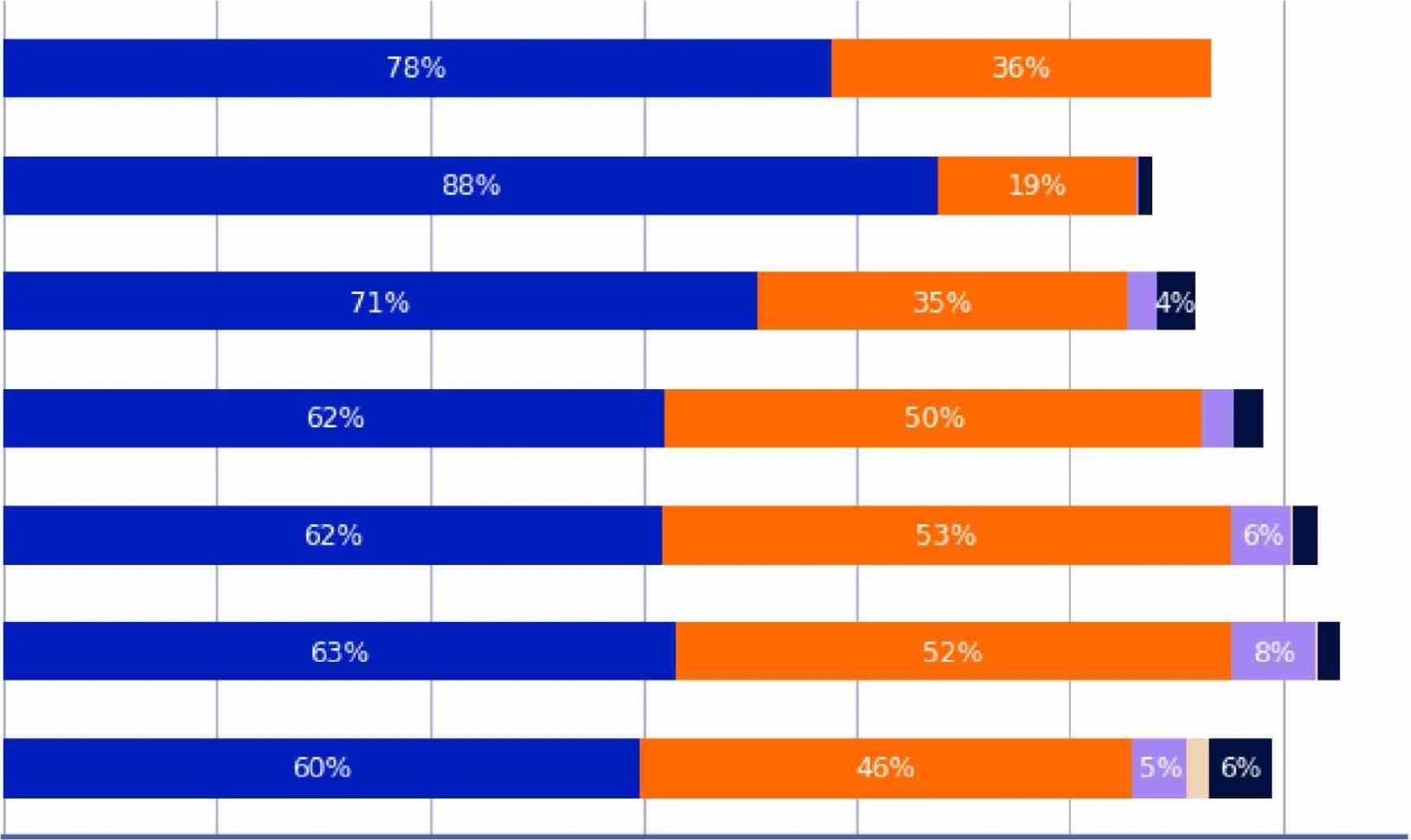
Percentage

Figure 20: Distribution of Abuse Type by Decade

### 7.4.1. Decade of Abuse by Ethnicity

#### Distribution of Ethnicity by Decade of Abuse

 Päkehä Mäori MELAA Pasifika Asian Other

Pre 1950

1950s

1960s

1970s

1980s

1990s

Post 2000s

0% 20% 40% 60% 80% 100% 120%

##### Percentage

Figure 21: Distribution of Ethnicity by Decade of Abuse

Note: Proportions add to greater than 100% as survivors can be attributed to more than one ethnicity.

### 7.4.2. Abuse Type by Age

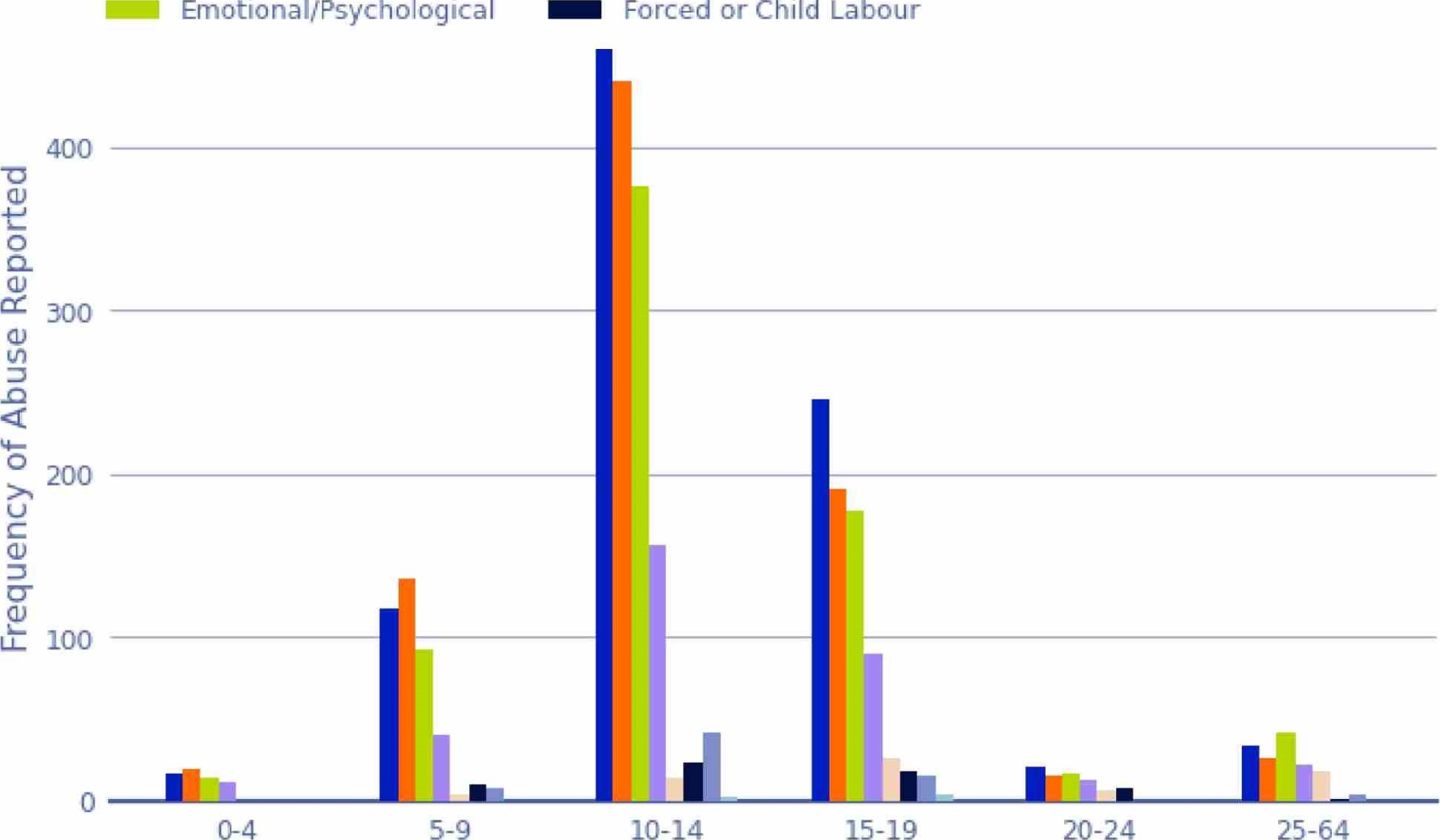
In this analysis we look at the frequency of abuse reported by survivors by their age and abuse type. This is calculated from the survivor time in care data, which records the age a survivor entered and left an institution, along with the instances of abuse reported during their stay. It is important to note that if a survivor reported abuse at a given institution between the ages of 5-12, it will be counted in both the 5-9 and 10-14 age groups. In cases where either the starting or ending age was absent from the data, instances of abuse were solely recorded for the available age category.

The chart below shows reported cases of sexual and physical abuse are relatively consistent among survivors in care within the 0-14 age bracket, with slightly elevated rates of sexual abuse. Among those aged 15 and above, instances of physical abuse were reported more frequently, For survivors aged 25+1 emotional abuse emerged as the most reported abuse type.

#### Frequency of Abuse Reported by Abuse Type and Age of Survivor

 Physical Neglect Sexual (non-contact)

 Sexual Medical Solitary Confinement



Emotional/Psychological

Forced

or

Child

Labour

Age of Survivor (at time of abuse)

Figure 22: Frequency of Abuse Reported by Type of Abuse and Age of Survivor

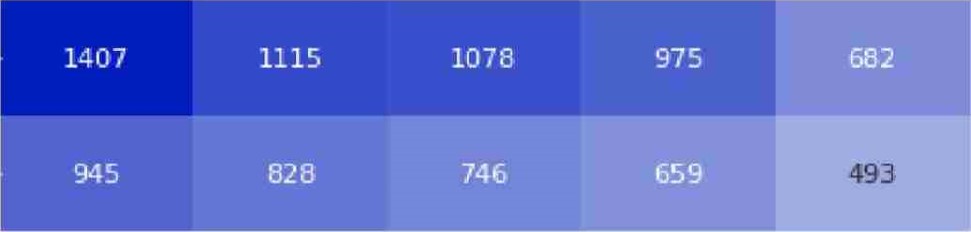
## 7.5. Abuse Type by Ethnicity

In our analysis of abuse types based on ethnicity, we've combined the Asian, MEL AA and Other survivor groups into the 'Other' category due to their relatively smaller population size. This categorisation results in the following population sizes for each ethnic group:

* Päkehä: 1,483
* Mäori: 1,018
* Pasifika: 113
* Other: 66

In the chart below, Päkehä experienced higher counts for each of the abuse types, which is relative to Päkehä making up 64% of the survivor population.

### Abuse Reporting by Ethnicity

 Päkehå194 150 132

Maori105 70 100

Pasifika 107 100 83 63 11 7

Other 66 55 55 36 11 9 8

Any Abuse Physical Sexual Emotional Neglect Medical Non-Contact Solitary

Sexual

#### Abuse Type



200 400 600 800 1000 1200 1400

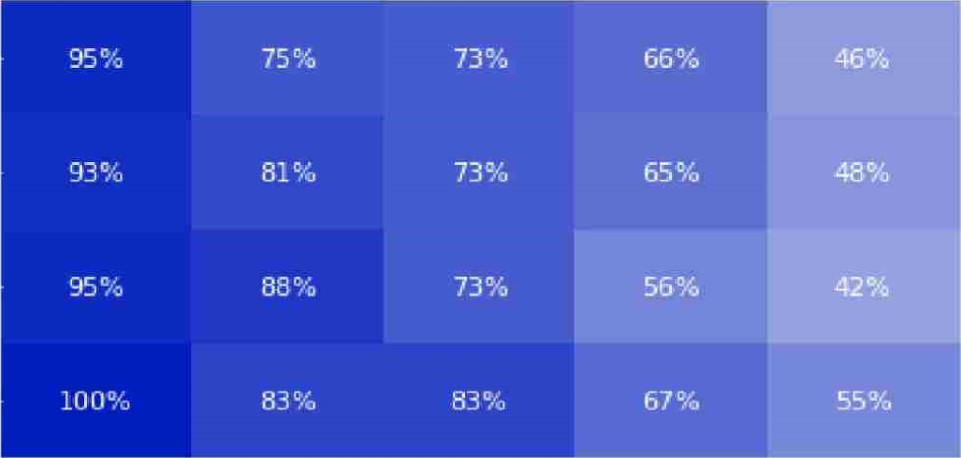
Survivor Abuse Counts

Figure 23: Abuse Reporting by Ethnicity

#### 7.5.1. Proportions by Ethnicity

When looking at the proportion of abuse types by ethnicity, Päkehä and Pasifika experienced the highest proportions of any type of abuse. Pasifika and Mäori experienced the highest proportion of physical abuse at 88% and 81% respectively. 'Other ethnicities' had the highest proportion of sexual abuse, and Päkehä had the highest proportion of emotional abuse.

### Distribution of Abuse Reporting by Ethnicity



Päkehå

13%

Mäori

10%

Pasifika

10%

Other

17%

14%

12%

Any Abuse Physical Sexual Emotional Neglect Medical Non-Contact Solitary

Sexual

Abuse Type

100%

Percentage of Survivors

Figure 24: Distribution of Abuse Reporting by Ethnicity

Below we conducted a series of proportion tests to determine whether particular ethnic groups exhibited a higher likelihood of experiencing specific types of abuse compared to the broader survivor population. Our attention was particularly focused on identifying differences within the Mäorj, Pasifika, and Päkehä survivor groups.

7.5.1.1. Mäori more likely to report physical abuse

Of the 1,018 Mäori survivors, 828 (81%) reported being physically abused compared to the non-Mäori survivor population, where 941 out of 1311 (72%) survivors reported physical abuse.

7.5.1.2. Pasifika more likely to report physical abuse

Of the 113 Pasifika survivors, 100 (88%) reported being physically abused compared to the non-Pasifika survivor population, where 1,669 out of 2216 (75%) survivors reported physical abuse. The proportions test of these findings indicates that the proportion of Pasifika people that reported physical abuse was larger than what would be expected due to chance.

7.5.1.3. Other significant relationships between ethnicity and abuse type:

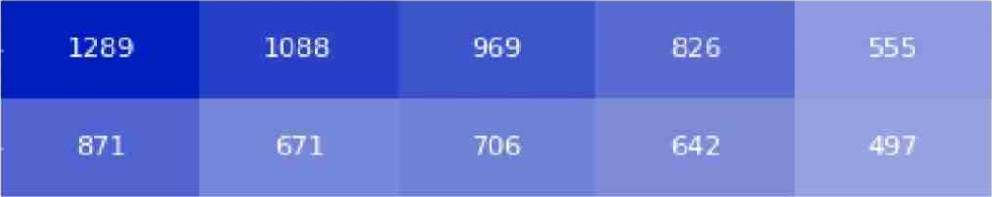
* Mäori survivors were more likely to experience neglect (48%, 493/1018, 95% Cl [45.4%, 51.5%]) compared to non-Mäori (43%, 567/1311, 95% a
* Päkehä survivors were more likely to experience emotional abuse (66%, 975/1483, 95% Cl [63.3%, 68.2%)] compared to non-Päkehä survivors (60%, 507/846 95% Cl [56.6%, 63.2%]).
* Päkehä survivors were more likely to experience medical abuse (13%, 194/1483, 95% a [11.4%/ 14.8%]) compared to non-Päkehä (10%, 85/846, 95% a \_[8.0%1 12.1%])
* Päkehä survivors were more likely to experience non-contact sexual abuse (10%,

150/1483, 95% Cl [8.5%, 11.6%]) compared to non-Päkehä (6%, 51/846, 95% CI [4.4%,

## 7.6. Abuse Type by Gender

When analysing abuse types by gender, males experienced a higher count across all abuse types, which is due to 59% of the survivor population being male.

### Abuse Reporting by Gender

 Male142 142 114

Female137 70 87

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Other | 17 | 10 | 10 | 14 |  |  |  |  |
|  | Any Abuse | Physical | Sexual | Emotional | Neglect | P&dical | Solitary | Non-Contact |

Sexual

Abuse Type



200 400 600 800 1000 1200

#### Survivor Abuse Counts

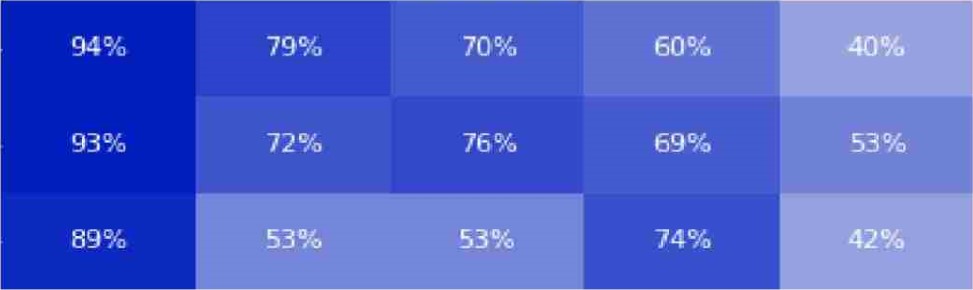
Figure 25: Abuse Reporting by Gender

Note: Due to sma// population sizes for gender groups other than male and female, these have been combined into "Other'( This includes: gender diverse, non binary, prefer not to say and not stated.

#### 7.6.1. Proportion of Abuse Type by Gender

When looking into the proportion of abuse types by gender, males are more likely to experience physical abuse and solitary abuse, whilst females are more likely to experience sexual, non-contact sexual and medical abuse.

#### Distribution of Abuse Reporting by Gender

Male8%

Q)

c Female15%

Other

Any Abuse Physical Sexual Emotional Neglect Medical Solitary Non-Contact

Sexual

Abuse Type

10% 20%

Percentage of Survivors

Figure 26: Distribution of Abuse Reporting by Gender

##### 7.6.2. Males more likely to report Physical Abuse

1,088 (79%) of the 1,378 male survivors reported being physically abused, while 681 non-male survivors out of 951 (72%) reported physical abuse.

This difference was statistica//y significant z = 4.01, p < .001, which indicates that the proportion of males that reported physical abuse was larger than what would be expected due to chance (95% C/  for male survivors, and 95% C/ [68.7%, 74.5%] for non-males).

##### 7.6.3. Females more likely to report Sexual Abuse

706 (76%) of the 932 female survivors reported being sexually abused, while 979 non-female survivors out of 1,397 (70%) reported sexual abuse.

This difference was statistica//y significant z = 4.01, p = .001, which indicates that the proportion of females that reported sexual abuse was larger than what would be expected due to chance (95% C/ [729%, 78.5%] for female survivors, and 95% C/ [67.6%, 72.5%] for non-females).

7.6.4. Other Significant Relationships between Gender and Abuse Type:

* Female survivors are more likely to report neglect (53%, 497/932, 95% Cl [50.1%, 56.5%]) compared to non-females (40%, 563/1,397, 95% Cl [317%, 42.9%]).
* Female survivors are more likely to report emotional abuse (69%, 642/932, 95% Cl [65.9%, 71.9%] compared to non-females (60%, 840/1,397, 95% a [57.6%, 62.7%]).
* Female survivors are more likely to report medical abuse (15%, 137/932, 95% Cl [12.4%, 17.0%]) compared to non-females (10%, 142/1,397, 95% Cl [8.6%, 11.7%]).  Male survivors are more likely to report solitary abuse (6%, 79/1378, 95% Cl [4.5%,

7.0%]) compared to non-males (4%, 35/951, 95% Cl [2.5%, 4.8%]).

## 7.7. Abuse Type by Ethnicity and Gender

### 7.7.1. Physical Abuse by Ethnicity and Gender

We reported earlier that Mäori and Pasifika had higher levels of physical abuse than the rest of the population. When analysing this alongside gender, we see that for all ethnicities, men had a higher percentage of physical abuse.

### Physical Abuse Reporting by Gender and Ethnicity

|  |  |  |  |
| --- | --- | --- | --- |
| 72% | 77% | 88% |  |
|  |  |  |  |
|  |

Female

Päkehä Mäori Pasifika Other

Ethnicity



72.5% 75.0% 77.5% 80.0% 82.5% 85.0% 81.5%

Percentage of Survivors

Figure 27: Distribution of Physical Abuse Reported by Gender

#### 7.7.2. Sexual Abuse by Ethnicity and Gender

We reported earlier a significant difference in the rates of abuse between men and women.

This relationship is more pronounced for Mäori, Pasifika and Other ethnicities than for Päkehä. 81% of Pasifika women reported sexual abuse, compared to 70% for Pasifika men. Mäori women reported a higher rate of sexual abuse compared to Mäori men. This relationship does not hold for Päkehä men and women, who reported sexual abuse at similar rates.

### Sexual Abuse Reporting by Gender and Ethnicity

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | | 74% | | |  |  |  | | --- | --- | --- | | 78% | 81% | 94% | |

Female

Male 72% 70% 74% 

Päkehä Mäori Pasifika Other

#### Ethnicity



70.0% 75.0% 80.0% 85.0% 90.0%

Percentage of Survivors

Figure 28: Distribution of Sexual Abuse Reporting by Gender

#### 7.7.3. Emotional Abuse by Ethnicity and Gender

When analysing the reports of emotional abuse by ethnicity and gender, females across all ethnicities report higher rates of emotional abuse, except for Pasifika, where the reporting of emotional abuse is 56% across both genders.

### Emotional Abuse Reporting by Gender and Ethnicity

|  |  |  |
| --- | --- | --- |
| Female  56%  56% | |  | | --- | | 71% | |

Pä kehä Mäori Pasifika Other

#### Ethnicity



56.0% 58.0% 60.0% 62.0% 64.0% 66.0% 68.0% 70.0%

Percentage of Survivors

Figure 29: Distribution of Emotiona/ Abuse Reporting by Gender

#### 7.7.4. Neglect by Ethnicity and Gender

Neglect is higher in females of all ethnicities. This was especially true for Pasifika, with Pasifika females reporting a 24% higher rate than Pasifika males,

## Neglect Reporting by Gender and Ethnicity

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  | | --- | --- | --- | | 53% | 59% | 65% | | 39% |  |  | |

Päkehä Mäori Pasifika Other

### Ethnicity



40.0% 45.0% 50.0% 55.0% 60.0%

Percentage of Survivors

Figure 30: Distribution of Neglect Reporting by Gender and Ethnicity

## 7.8. Abuse Type by Mental Distress

The proportion of survivors who reported abuse is higher for all abuse types for people who have also experienced mental distress at some point. This difference is greatest for survivors reporting sexual abuse, where survivors who reported a mental distress condition had a 19% higher reporting rate of sexual abuse.

### Abuse Type by Mental Distress

|  |  |  |  |
| --- | --- | --- | --- |
|  | 48% | 13%  5% 7% | 95%  87% |
|  |

Unidentified

Physical Sexual Emotional Neglect Bkdical Non-Contact Any Abuse

Sexual

#### Abuse Type

10% 

Percentage of Survivors

Figure 31: Distribution of Abuse Type by Mental Distress

## 7.9. Abuse Type by Sexual Identity

Abuse reporting rates are significantly higher for LGBTQIA+ members for sexual, emotional and non-contact sexual abuse.

### Abuse Type by Sexual Identity

 98%

|  |  |  |
| --- | --- | --- |
|  | 46%  46% | 12% 14%  12% 8% |
| 63% |

No/lJnidentified93%

Physical Sexual Emotional Neglect Medical Non-Contact Any Abuse

Sexual

#### Abuse Type

20%

Percentage of Survivors

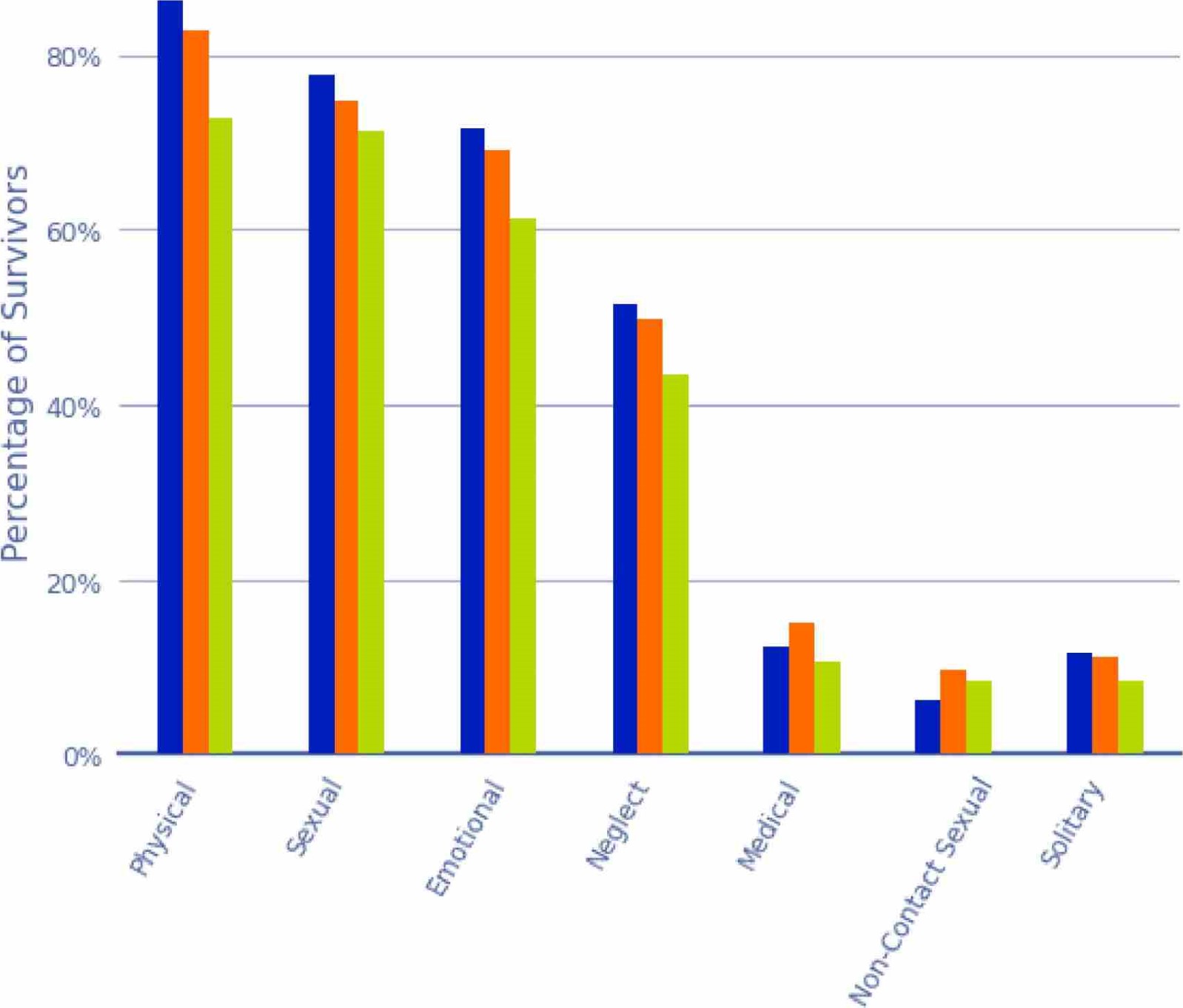
Figure 32: Distribution of Abuse Type by Sexual Identity

## 7.10. Abuse Type by Deaf or Disability

In the chart below we've analysed the prevalence of abuse reported by Deaf or disabled survivors. The data shows that survivors who are Deaf or disabled typically report abuse types at a higher rate. There are 624 survivors with a disability, 130 who are Deaf, and 1,625 who are neither Deaf nor disabled

Abuse by Deaf or Disability

 Deaf Survivors with a disability Survivors not Deaf or disabled



### Abuse Type

Figure 33: Distribution of Abuse Type by Deaf or Disability

Note that 50 Deaf survivors a/so reported a disability and are counted in both categories.

Using a two proportion z-test, we found that Deaf survivors are more likely to report physical

.001, z=3.3) and emotional abuse (p=0.0191 z=2.3) than survivors who are not Deaf or disabled.

Because of larger sample sizes, more significant differences exist between the disabled and not Deaf or disabled groups, with the disabled group consistently reporting higher rates of abuse.

* Disabled survivors are more likely to report physical abuse (13% p < .001, z=5.0) compared with survivors who are not Deaf or disabled.  Disabled survivors are more likely to report emotional abuse (13% p < .001, z=3.6) compared with survivors who are not Deaf or disabled.  Disabled survivors are more likely to report medical abuse (40% p=0.0043, z=2.9) compared with survivors who are not Deaf or disabled.
* Disabled survivors are more likely to report neglect (14% p=0.0067, z=2.7) compared with survivors who are not Deaf or disabled.
* Disabled survivors are more likely to report solitary abuse (33% p=0.042, z=2.0) compared with survivors who are not Deaf or disabled.

## 7.11. Abuse Type by Ethnicity and Deaf or Disability

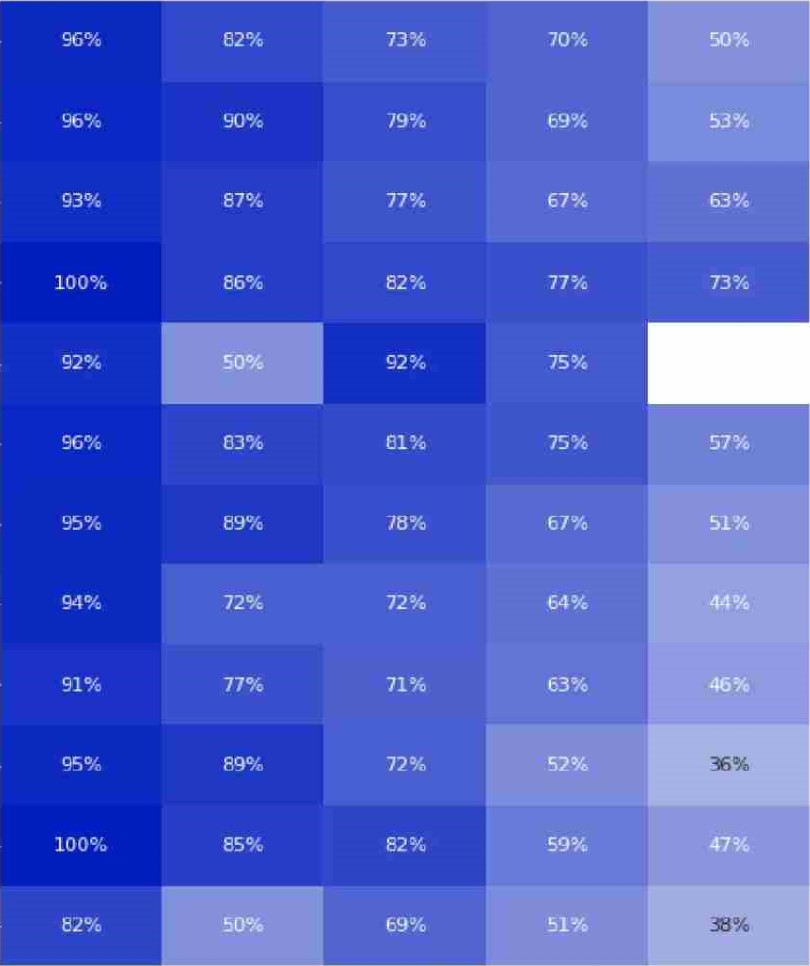
Previously, we found that the proportion of survivors who are Deaf or living with a disability report abuse at a significantly higher rate than survivors who are not. This trend was consistent across all forms of abuse, with the exception of non-contact sexual abuse, which did not reach statistical significance at the 5% level

To delve deeper, this section explores the intersection of those who are Deaf or living with a disability and ethnicity to determine whether there are specific ethnic groups which have a higher prevalence of abuse among survivors with disabilities.

Importantly we note that Mäori survivors with a disability report the highest prevalence of physical abuse at (90%) followed by Mäori survivors who are Deaf (89%).

The chart below summarises reporting rates of each abuse type for every combination of disability group and ethnicity.

Abuse Type by Deaf or Disability and Ethnicity

 Survivors with a disability-Päkehä15% 11% 11%

Survivors with a disability-Mäori

Survivors with a disability-Pasifika

Survivors with a disability-Other

Survivors With a disability-Umdentified

Deaf-Päkehä13% 12%

æaf-Mäori11% 10% o

 Survivors not Deaf or disabled-Päkehä 12% 10% 

Survivors not Deaf or disabled-Mäori6% 10%

Survivors not Deaf or disabled-Pasifika

Survivors not Deaf or disabled-other18%

Survivors not Deaf or disabled-unidentified

Any Abuse Physical Sexual amotional Neglect P&dical Non-Contact Solitary

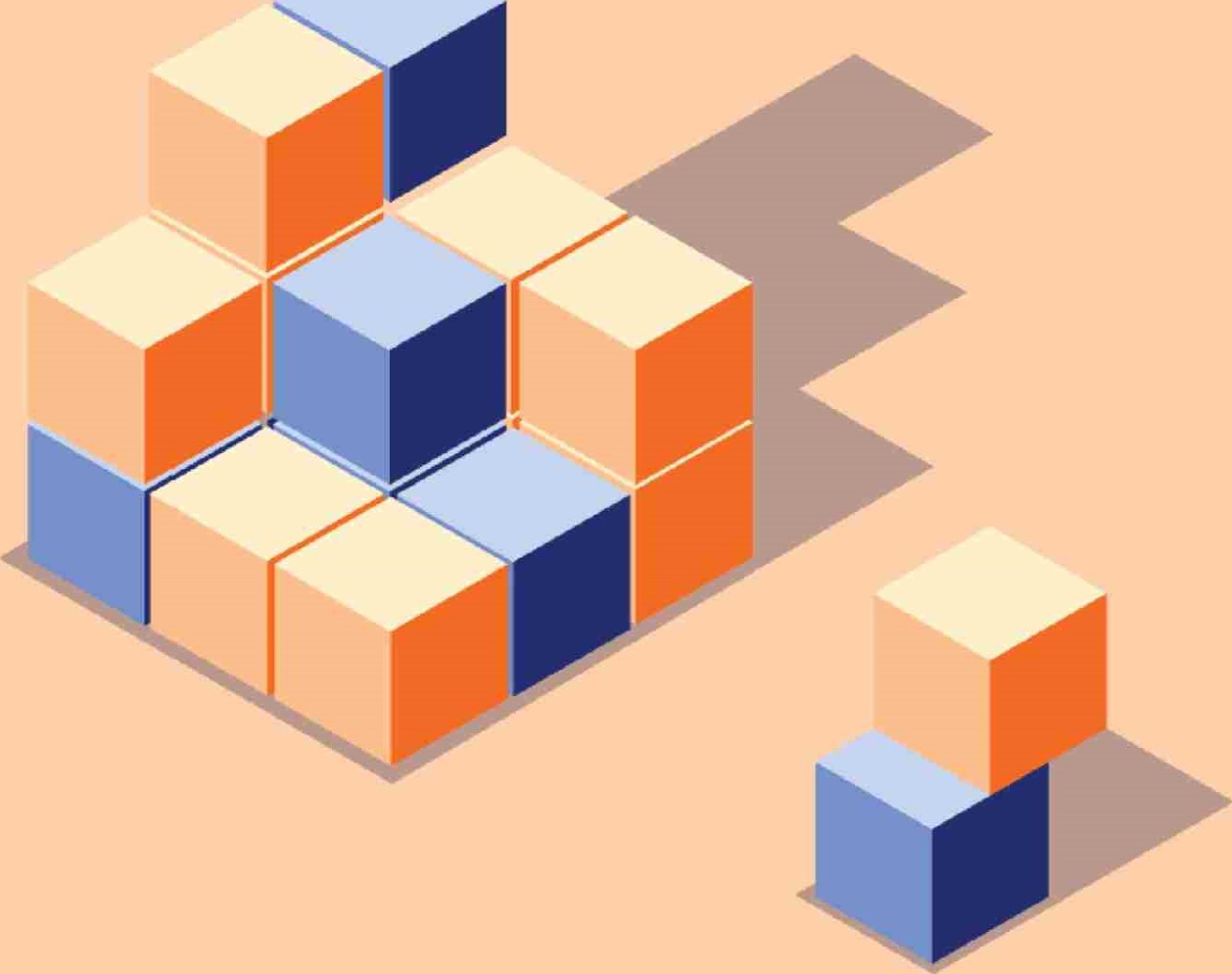
20% 100%

Percentage of Survivors

Figure 34: Abuse Types by Deaf and Disability and Ethnicity



Institution and Setting



8. Institution and Setting

In this section we analyse the institutions, institution types and settings that provided care for survivors with accounts, noting that not all survivors knew the institutions that they went to or could recall the names of the institutions.

### 8.1. Count by Institution

The institutions with the highest number of unique survivors are listed in the table below:

|  |  |
| --- | --- |
| Institution | Unique Survivors |
| Foster Care | 607 |
| Adoption | 168 |
| Epuni Boys Home | 144 |
| Family Homes | 131 |
| Owairaka Boys Home | 126 |

Table 1: Counts of Unique Survivors by Institution (for the five most attended Institutions)

## 8.2. Count by Setting

The table below shows the number of unique survivors by setting:

|  |  |  |
| --- | --- | --- |
| Setting Type | Setting | Unique  Survivors |
| Social Welfare and Youth Justice | Total | 1338 |
| Borstal | 159 |
| Boys' or Girls' home | 766 |
| Care Provider | 218 |
| Foster care, foster homes, family homes | 715 |
| SW Government Department | 217 |
| Faith | Total | 815 |
| Church | 148 |
| Faith Community | 80 |
| Faith-based orphanages, residences and training centers | 278 |
|  | Foster/family homes | 21 |
| Religious schools | 249 |
| Service Provider | 125 |
| Unmarried Mothers home | 21 |
| Disability and Mental Distress | Total | 437 |
| Disability Care Setting | 29 |
| Forensic Care | 8 |
| Psychiatric Care | 321 |
| Special School | 105 |
| Supported Living | 20 |
| Other | Total | 301 |
| Government Department (non-SW) | 5 |
| Hospital | 141 |
| Other | 171 |
| Education | Total | 153 |
| Education | 153 |
| Transitional and Law enforcement | Total | 138 |
| Transitional and Law enforcement | 138 |
| Health Camps | Total | 51 |
| Health Camp | 51 |
| Deaf Schools | Total | 16 |
| Deaf Schools | 16 |
| Blind Schools | Total | 2 |
| Blind School | 2 |

Table 2: Counts of Unique Survivors by Setting

## 8.3. Setting by Length of Time in Care

This graph uses imputed records to establish the length of time in care. 2/843 (49.9%) records out of 5,697 have been imputed based on setting, which reflects the total number of records of unique institutions that survivors attended.

The graph below details the length of time in care per setting. Faith Communities had the highest length of time in care, with up to 52 years in care

### Length of Time in Care 'by Setting

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |
|  |  |
|  |  |
|  |  |
|  |  |  |  |  |
|  |
|  |
| I |  |
|  |
|  |  |
|  |  |
|  |  |
|  |

Social Welfare and Youth Justice

Borstal Boys' or Girls' home Care Provider

Foster care, foster homes, family homes

SW Government Department

Faith Church

Faith Community

Faith-based orphanages, residences and training centers

Fosterffamily homes Religious schools Service Provider Unmarried Mothers home

Disability and Mental Distress Disability Care Setting

Forensic Care

Psychiatric Care Special School Supported Living

Other

Government Department (non-SW)

Hospital Other

Education Education

Transitional and Law enforcement

Transitional and Law enforcement

Health Camps Health Camp

Deaf Schools

Deaf School 

10 20 30 40 50

Years in Care

Figure 35: Length of Time in Care by Setting

Note the Blind School setting has been excluded due to incomplete data.

## 8.4. Setting by Experience of Incarceration

The table below shows the percentage of survivors who have been incarcerated at some point in their life, by setting. From this, we can derive that based on the survivors who were in forensic care, borstal, transitional and law enforcement or girls or boys home settings, at least 50% have been incarcerated at some point in their lives.

Of the 159 survivors who have been in borstal settings, 100 or 63% of those have been incarcerated at some point during their life.

|  |  |  |  |
| --- | --- | --- | --- |
| Setting Type | Setting | Incarcerated Survivors by Setting | Percentage of All Survivors |
| Social Welfare and Youth Justice | Total | 548 | 41% |
| Borstal | 100 | 63% |
| Boys' or Girls' home | 390 | 51 % |
| Care Provider | 84 | 39 % |
| Foster care, foster homes, family homes | 260 | 36% |
| SW Government Department | 70 | 32 % |
| Faith | Total | 166 | 20% |
| Church | 15 | 10 % |
| Faith-based orphanages, residences and training centers | 90 | 32 % |
| Foster/family homes | 6 | 29% |
| Religious schools | 44 | 18 % |
| Service Provider | 35 | 28 % |
| Disability and  Mental Distress | Total | 150 | 34% |
| Disability Care Setting | 6 | 21 % |
| Forensic Care | 6 | 75% |
| Psychiatric Care | 108 | 34 % |
| Special School | 47 | 45 % |
| Other | Total | 82 | 27% |
| Hospital | 37 | 26 % |
|  | Other | 51 | 30% |
| Education | Total | 49 | 32% |
| Education | 49 | 32 % |
| Transitional and  Law enforcement | Total | 82 | 59% |
| Transitional and Law enforcement | 82 | 59% |
| Health Camps | Total | 25 | 49% |
| Health Camp | 25 |  |

Table 3: Incarcerated Survivors by Setting

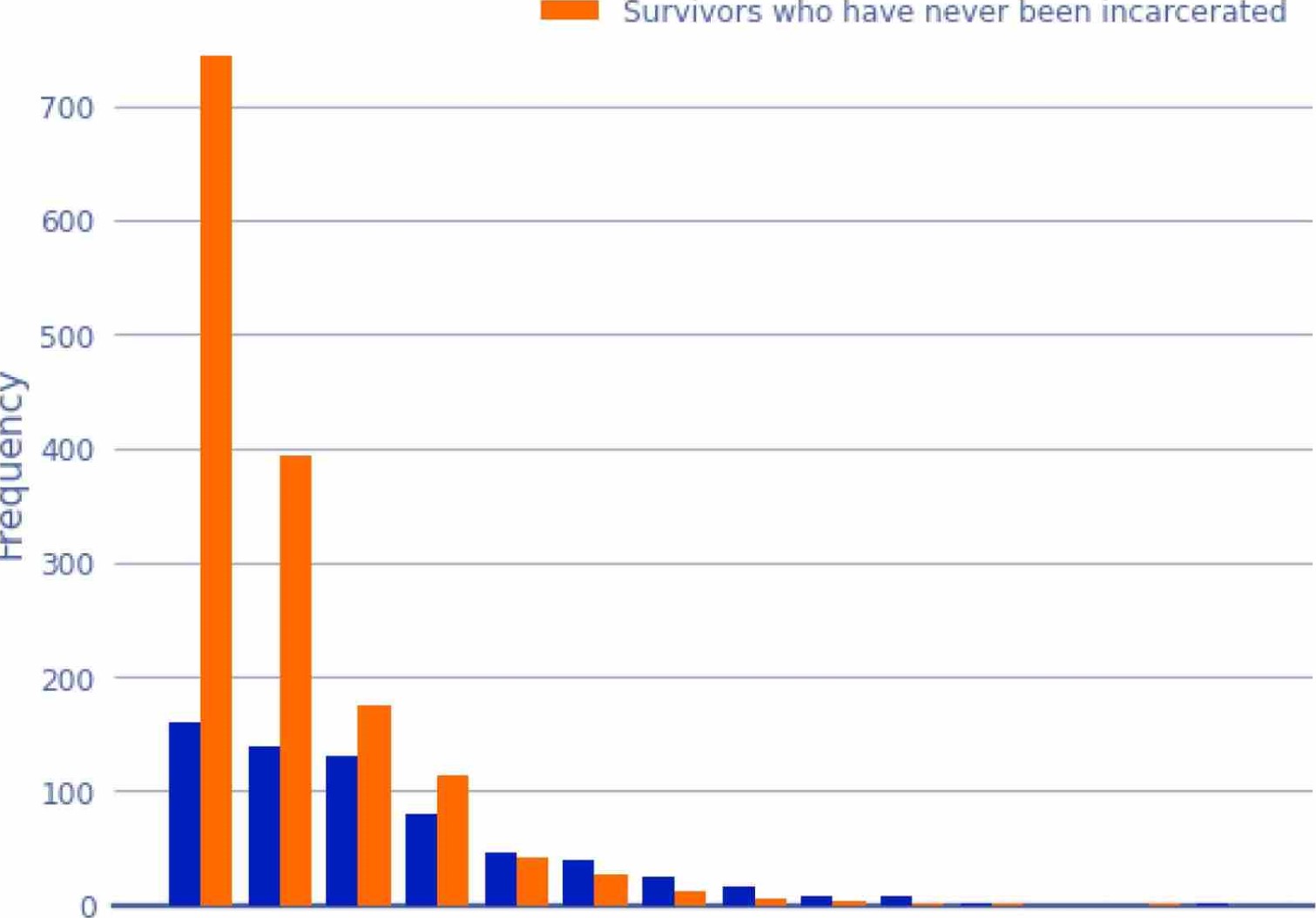
Some settings in the tab/e above have been suppressed to maintain confidentiality.

### 8.5. Count of Institutions by Incarceration

In the graph below we look at the frequency of institutions attended by incarceration rate, The analysis shows survivors were much more likely to be incarcerated if they attended five or more institutions, compared with those survivors who were placed in one to four institutions.

## Unique Institutions by Incarceration

 Survivors who have been incarcerated

Survivors who have never been incarcerated

1 2 3 4 5 6 7 8 9 10 11 12 13 14

Unique Institutions Visited

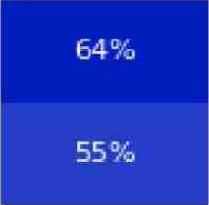
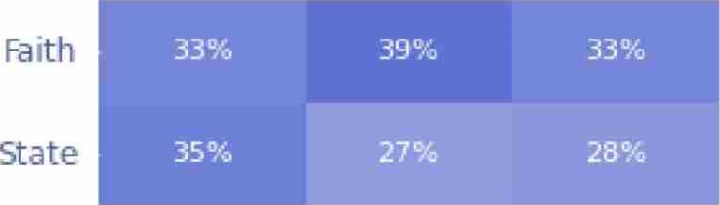
Figure 36: Unique Institutions by Incarceration

## 8.6. Institution Types by Abuse Type

When comparing state institutions with faith-based institutions, we can see that the rates of physical abuse are similar across both of these institution types. Sexual and emotional abuse is more prevalent in faith-based institutions.

### Abuse Type by Institution Type

|  |
| --- |
| 15% |

 1% 11% 4%

3%2%

Physical Sexual Emotional t&dical Neglect Non-Contact Solitary Any Abuse

Sexual



Percentage of Survivors

Figure 37: Institution Types by Type ofAbuse

## 8.7. Setting by Abuse Type

Emotional abuse was highly prevalent at unmarried mothers' homes. 70% of survivors who resided at this setting reported emotional abuse. This is the same proportion as the 'Any Abuse" category.

The church setting has the highest proportion of sexual abuse reported at 50%, but a comparatively lower proportion of physical abuse at 20%. The Psychiatric setting has the highest reported medical abuse at 11%.

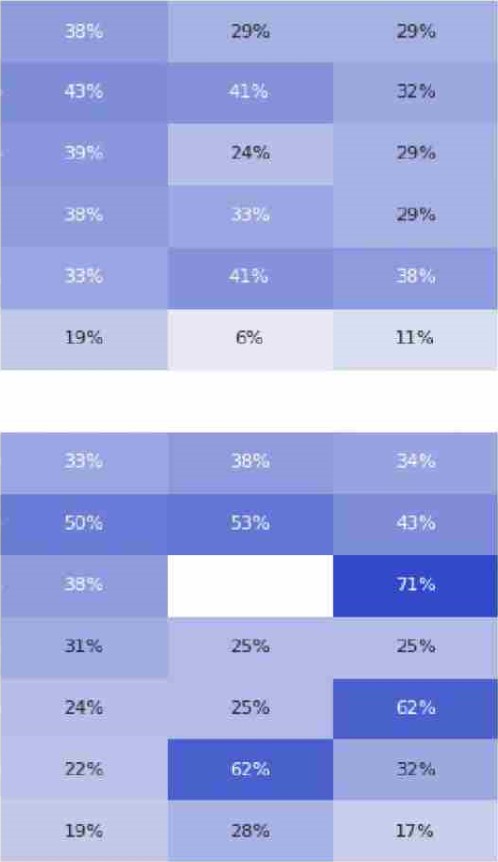
Abuse Type by Setting

social welfare and youth Justice 

|  |
| --- |
|  |
| 76% |
|  |

|  |  |
| --- | --- |
| 63% | |
|  |  |
|  | |
|  | |
|  | |
|  | |

Foster care, foster homes, family homes19%



1%

1%

1%

1%

Boys' or Girls' horne12%

Care Provider

SW Government Department31%

Borstal

Faith4%

Religious schools

unmarried Mothers home 

Faith-based orphanages.4% residences and training centers

Faith Community20%

Church

service provider

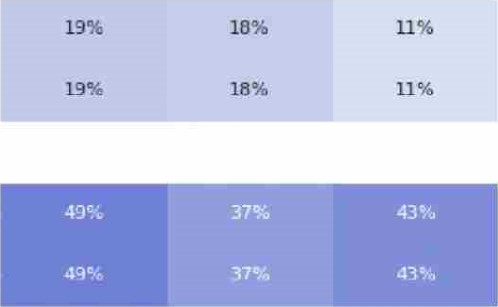
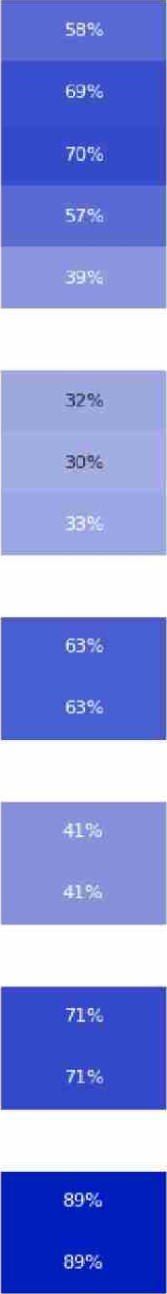
Foster,ifamily homes

|  |
| --- |
| 29% |

|  |
| --- |
| 20% |

|  |  |
| --- | --- |
| 29%  29% |  |

|  |  |  |
| --- | --- | --- |
|  |  | 44%  44% |

Disability and Mental Distress17%

Special school16%

Supported Living30%

Psychiatric Care11% 17% 1%

Disability Care Setting 19%

Other  4% 10%

Other 16% 14%

Hospital  12%

Education

Education

Transitional and Law enforcement

Transitional and Law enforcement

Health Camps14%

Health Camp14%

Deaf Schools

Deaf School

Physical Sexual Emotional Neglect Non-contact Any Abuse

Sexual



10% 20% 30% 40% 50% 60% 70% 80%

Percentage

•

Figure 38: Setting by Abuse Type

Note: Settings that had fewer than 20 survivors in care have been removed from this analysis.

### 8.8. Institution by Abuse Type

When analysing the institutions with the highest reported accounts of the types of abuse, we found that School had the highest proportion of survivors report sexual abuse at 75%. had the fourth highest rate of reporting sexual abuse at 59%.

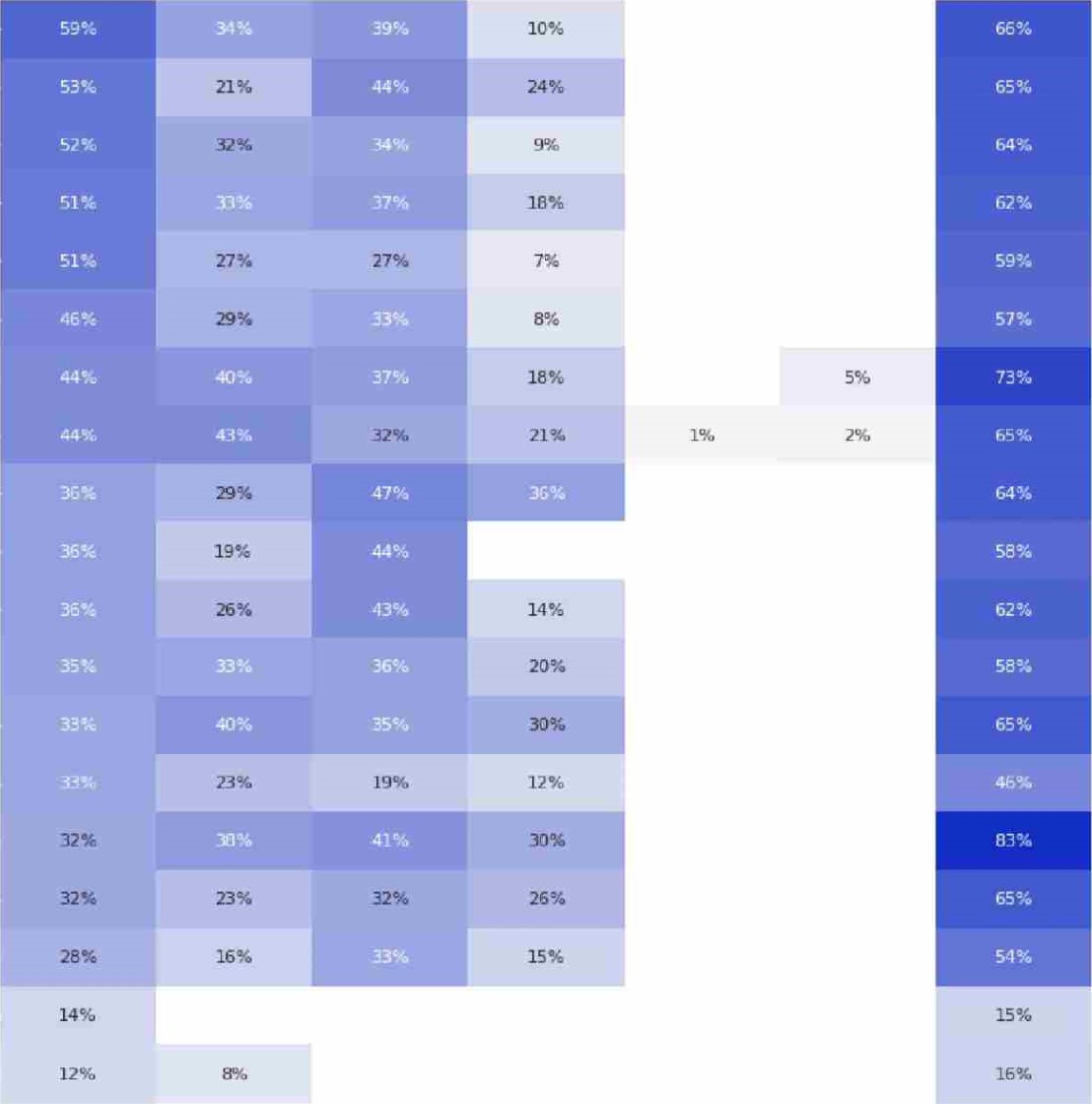
The highest level of physical abuse was reported at Wesleydale Boys Home and Owairaka Boys Home (also known as Auckland Boys Home). 68% of Wesleydale Boys Home and 59% of Owairaka Boys Home survivors reported being physically abused.

The had the highest reporting rate of emotional abuse at 85%.

Abuse Type by Institution

social Welfare and Youth Justice

Boys Home 



Owairaka

Boys

Home

Christchurch

Boys'

Horne

Kohitere

Boys

Training

Centre

Epuni

Boys

Home

Hokio

Beach

School

Hamilton

Boys

Home

Family

Homes

Foster

Care

Bollard

Girls

Home

Dunedin

Boys

Home

Miramar

Girls

Horne

State

Ward

Child,

youth

and

Family

services

(1999-2017)

Weymouth

Residential

centre

Department

of

Social

Welfare

1972-1992

Christchurch

Girls'

Home

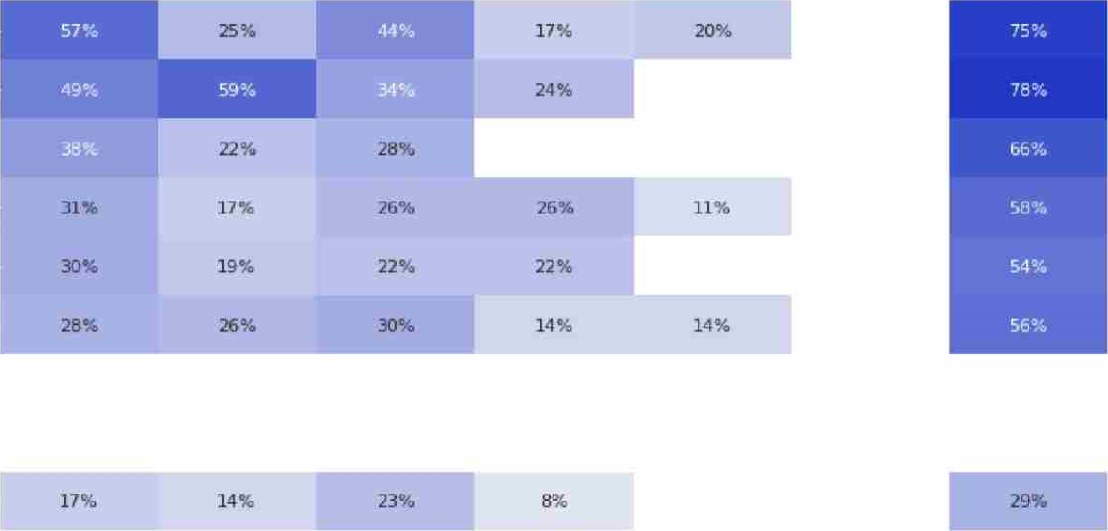
Kngslea

Corrective

Training

Borstal

Disability and Mental Distress



Lake

Alice

Hospital

Marylands

School

Cherry

Farm

Porirua

psychiatric

Hospital

Sunnyside

Tokanui

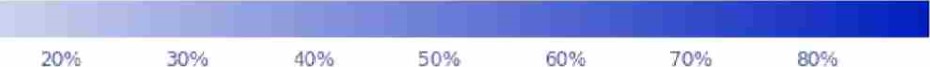
Psychiatric

Hospital

Other

Adoption

mysical Sexual Emotional Neglect '&dical Non-Contact Any Abuse sexual

10% 

Percentage

•

Figure 39: Institution by Type of Abuse

Note: Institutions with mu/tip/e names have been reduced to one on the graph above e.g. Hamilton Girls Home/Dey Street Residence has been renamed as Hamilton Girls Home. Institutions that had fewer than 30 survivors reside there have been removed from this analysis.

8.9. Demographic Prevalence at Institutions and Settings

#### 8.9.1. Mental Distress

Of all survivor accounts, 83% of survivors reported living with mental distress at some stage during or following their state or faith-based care.

##### 8.9.2. Mental Distress by Institution type

State institutions have a higher proportion of survivors who reported mental distress at some point in time (note that this may be after leaving an institution).

* 84.8% of survivors at state institutions reported experiencing mental distress during or following their care experience (1460/1721)
* 82.0% of survivors at faith institutions reported experiencing mental distress during or following their care experience (691/843)

Considering only unique instances, where an individual was admitted to the same institution multiple times, the prevalence of mental distress is higher:

* 87.1% of instances at state institutions involve survivors who report experiencing mental distress (3952/4537)
* 83.6% of instances at faith institutions involve survivors who report experiencing mental distress (975/1166)

##### 8.9.3. Mental Distress by Setting

The prevalence of mental distress by setting is detailed in the following table. The table shows that the rate of mental distress reported by survivors due to admission at any setting is high, with a range from 100% at its highest to 62% at its lowest across all settings.



|  |  |  |  |
| --- | --- | --- | --- |
| Instances Percentage of All Setting Type Setting of mental distress Survivors | | | |
| Social Welfare and  Youth Justice | Total | 2728 | 86% |
| SW Government Department | 210 | 88% |
| Boys' or Girls' home | 1328 |  |
| Borstal | 173 | 86% |
| Foster care, foster homes, family homes | 817 | 86% |
| Care Provider | 200 | 82% |
| Faith | Total | 930 | 84% |
| Unmarried Mothers home | 20 | 95% |
| Foster/family homes | 22 | 92% |
| Faith-based orphanages, residences and training centers | 336 | 86% |
| Service Provider | 111 | 85 % |
| Religious schools | 250 | 82% |
| Church | 129 | 81 % |
| Faith Community | 62 |  |
| Disability and Mental Distress | Total | 590 | 90% |
| Forensic Care | 9 | 100% |
| Psychiatric Care | 433 | 92% |
| Supported Living | 18 | 90% |
| Special School | 100 | 85% |
| Disability Care Setting | 30 | 83 % |
| Other | Total | 317 |  |
| Government Department (non-SW) |  | 100% |
| Hospital | 152 |  |
| Other | 158 | 88% |
| Education | Total | 166 | 90% |
|  | Education | 166 | 90% |
| Health Camps | Total | 42 | 82% |
| Health Camp | 42 | 82% |
| Transitional and  Law enforcement | Total | 140 | 85% |
| Transitional and Law enforcement | 140 | 85% |
| Deaf Schools | Total | 13 | 62 % |
| Deaf Schools | 13 | 62% |
| Blind Schools | Total | ..c | ..c |

Table 4: Instances of Mental Distress by Setting

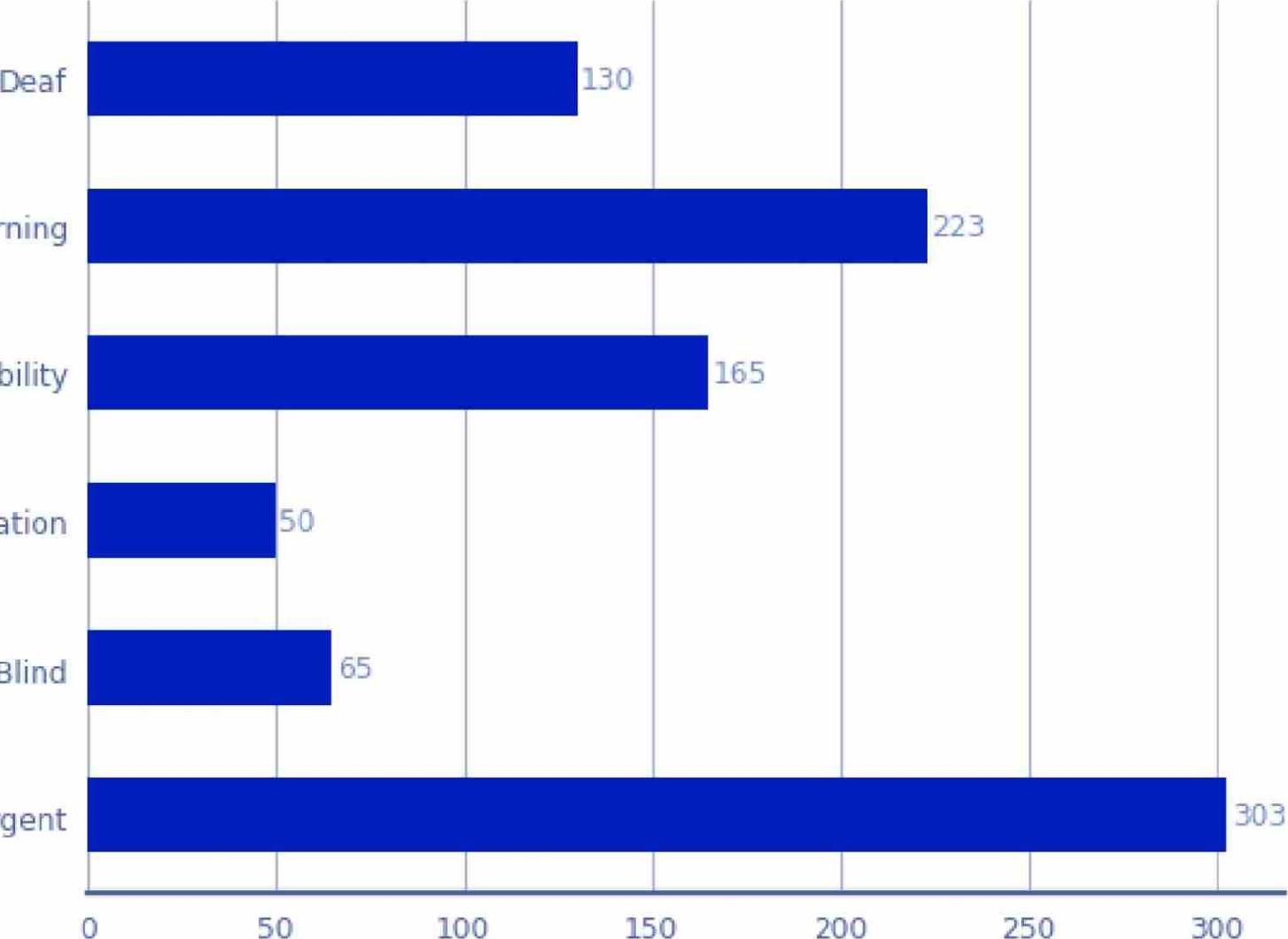
We have removed settings with fewer than 6 people. Note that this data is based on instances and not individuals. This means that a survivor who appeared at mu/tip/e institutions, or at the same institution mu/tip/e times wi// be counted twice.

##### 8.9.4. Deaf and Disability

Out of 2,329 survivors, 88% (2,045 people) reported being Deaf or living with a disability. The largest disability groupings were survivors with neurodiversity, learning disabilities and mobility disabilities.

Below we have detailed the breakdown of Deaf or disability types, noting that an individual may have multiple disabilities.

### Occurrence of Deaf or Disabilities

Disability - Learning

 Disability - Mobility

 Disability - Communication

Disability - Blind

Neurodivergent

Number of Survivors

Figure 40: Occurrence of Deaf or Disabilities

#### 8.10. Abuse Type by Deaf or Disability

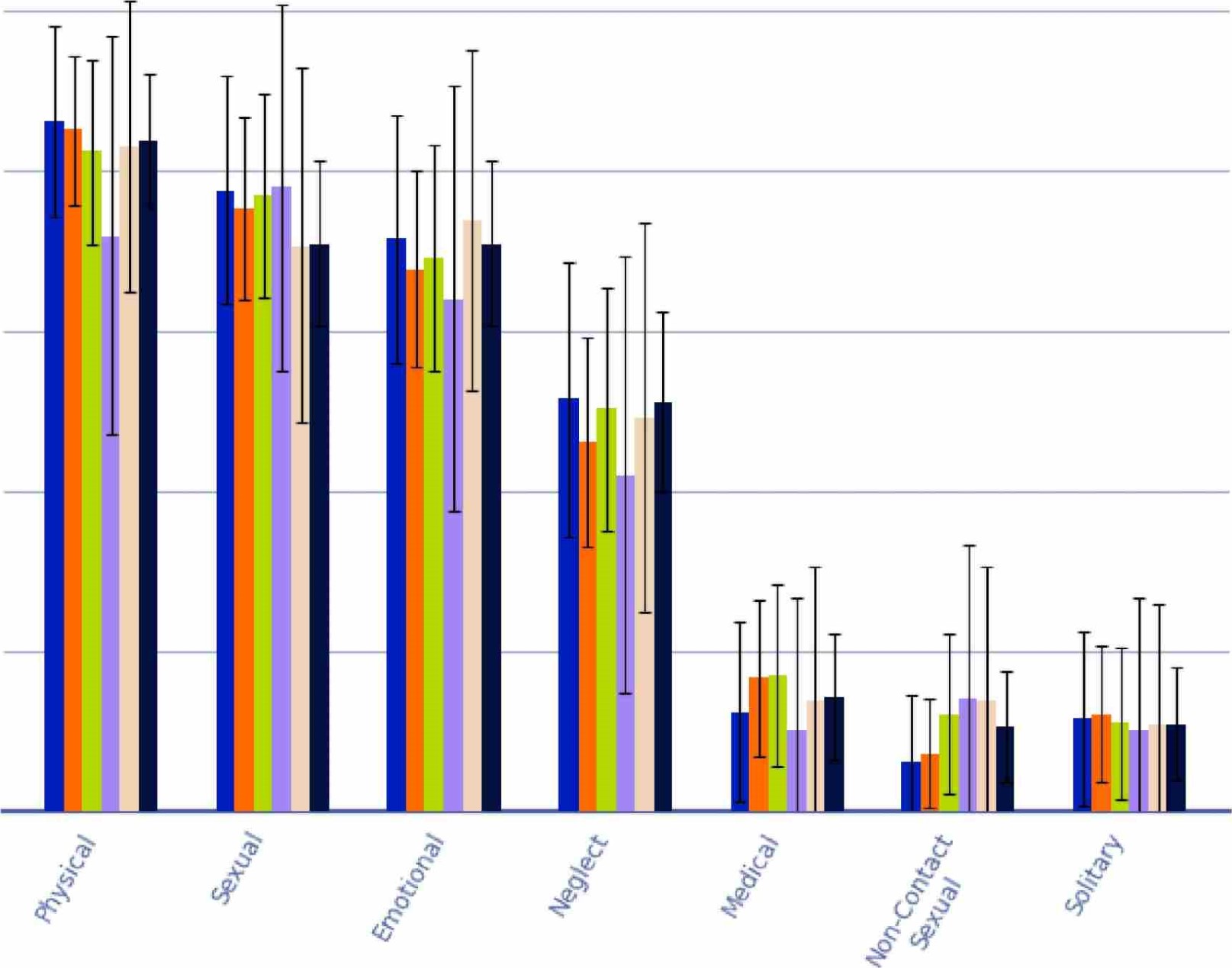
Plotted on the graph below are the rates of abuse by survivors who are Deaf or living with a disability. The 95% confidence intervals are indicated by the error bars. While many of these disabilities will relate to the time after being in care, we can still draw conclusions relating to whether or not being subject to a specific abuse type results in some disability later on.

Due to the high overlap between 95% confidence intervals in the graph above, we conclude that there is no difference between survivors living with a disability or who are Deaf groups as to the abuse they suffered.

### Frequency of Abuse by Deaf or Disability

 Deaf Disability - MobilityDisability - Blind

 Disability - Learning Disability - Communication  Neurodivergent

1.0

0.8

u 0.6

0.4

0.2

0.0

Abuse Type

Figure 41: Frequency of Abuse by Deaf or Disability

#### 8.10.1. Deaf or Disability Types by Setting Type

To determine if some setting types had higher levels of survivors living with a disability or who are Deaf, we analysed the prevalence of survivors living with a disability or who are Deaf at every institution. We have detailed the settings types with the highest ratio (based on setting types with more than 10 people). From this we found:

#### Highest ratio of Deaf

* 1.0 : Deaf Schools
* 0.11 : Education
* 0.078 : Health Camps

#### Highest ratio of Disability - Learning

* 0.19 : Disability and Mental Distress
* 0.16 : Education
* 0.14 : Health Camps

#### Highest ratio of Disability - Mobility

* 0.16 : Disability and Mental Distress
* 0.14 : Deaf Schools
* 0.11 : Other

#### Highest ratio of Disability - Communication

* 0.065 : Education
* 0.048 : Deaf Schools
* 0.04 : Other

#### Highest ratio of Disability - Blind

* 0.065 : Education

 0.048 : Deaf Schools

* 0.046 : Other

Highest ratio of Neurodivergent

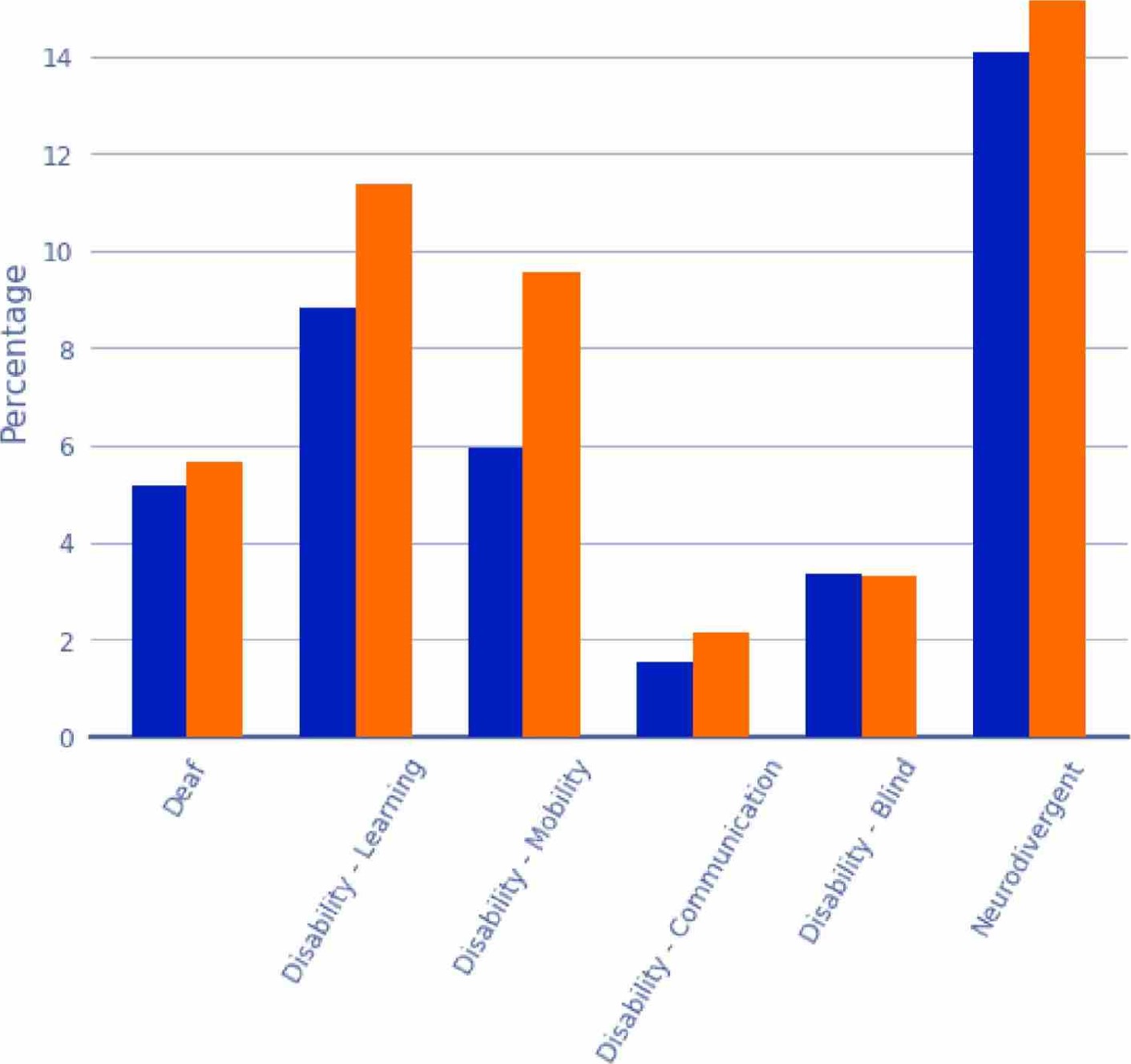
* 0.35 : Health Camps
* 0.21 : Transitional and Law enforcement
* 0.21 : Education

##### 8.10.2. Deaf or Disability by Institution Type

Survivors who are Deaf or living with a disability are more widely represented in state institutions across all categories, except for blindness, which has a 3.3% prevalence in both state and faith-based institutions.

### Deaf or Disability by Institution Type

 Faith State

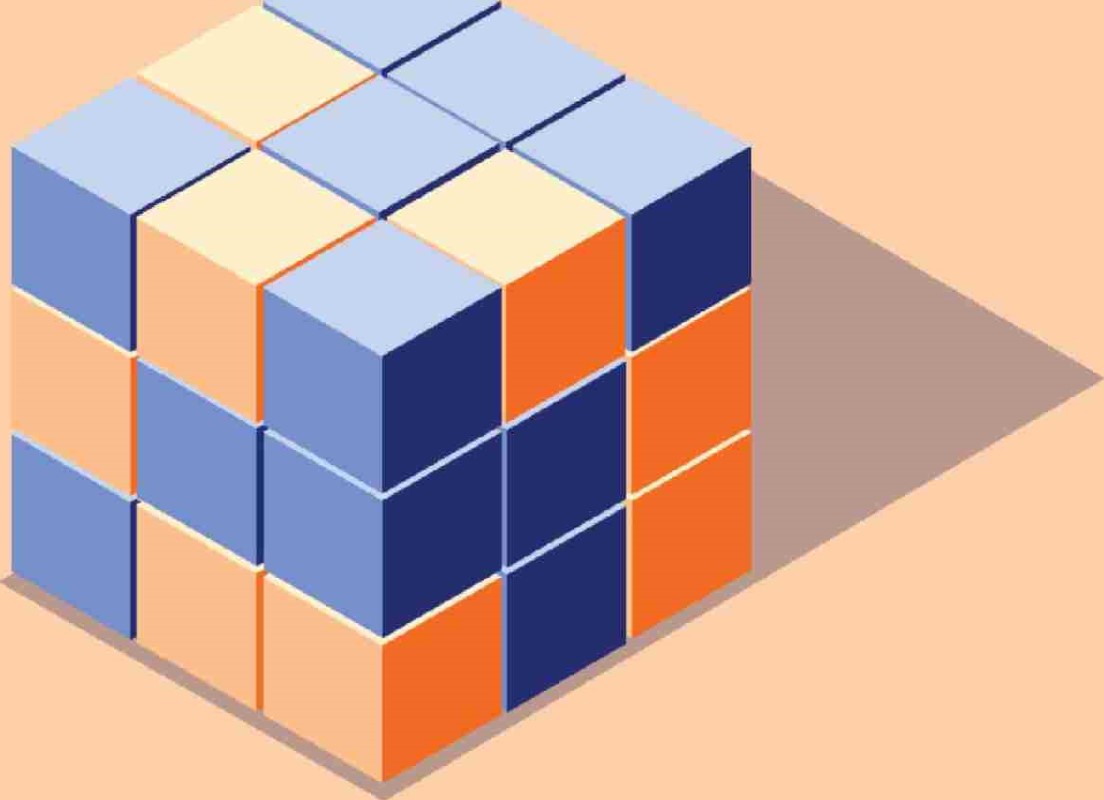


#### Deaf or Disability

Figure 42: Deaf or Disability by Institution Type



## Conclusion



### 9. Conclusion

In summary, the Abuse in Care Royal Commission of Inquiry has undertaken an exhaustive examination of the abuse reported by survivors of state and faith-based care systems in Aotearoa New Zealand between 1950 and 1999. Our quantitative analysis has augmented the Royal Commission's analysis by introducing an empirical dimension to the qualitative reporting.

Through a comprehensive investigation of 2,329 survivor narratives, our approach has detailed discernible trends and patterns within the broader tapestry of survivor experiences, including their pathway into care, the survivor care experience by setting, the abuse types suffered, and the lasting impacts on them as people.

Our analysis has confirmed the offending levels by time period, setting and demographic. The decade of the 1970s has emerged as a nexus of heightened abuse incidents, correlating with elevated institutional populations. The report outlines disparities in the occurrence of abuse types, particularly affecting pre-teens and teenagers, Mäori, Pasifika, and survivors who live with a disability or who are Deaf.

Young survivors aged 10 - 14 endured extreme levels of sexual and physical abuse. Mäori and Pasifika endured higher levels of physical abuse than other ethnicities and survivors with disabilities suffered higher levels of abuse across all abuse types. The analysis conclusively demonstrates that the age, ethnicity and disability level of survivors played a significant role in the abuse they were subjected to by caregivers.

The consequences of the abuse and neglect are also evident within the report, with 83% of survivors impacted by enduring mental distress. Moreover, the analysis showed that 29% of survivors were incarcerated at some point in their lives, and 10% have been involved in gang-related activities.

In conclusion, the data-driven analysis of the survivor accounts has confirmed the significant abuse and neglect reported by survivors as care-dependent individuals and makes clear the enduring sustained negative impacts the care experience has had on each survivors life.

10. Appendix

### 10.1. Definitions of Variables

|  |  |
| --- | --- |
| Variable | Definition of Variable |
| Age | Current age of survivor at time of registration. |
| Gender | What the person regards themself as at time of registration. |
| Country | The country where the survivor currently resides. |
| Region | The region where the survivor currently resides. |
| Ethnicity | As many ethnicities the person identifies with. |
| LGBTQIA+ | Yes if positively states - different terminology used - including lesbian, gay, bisexual, transgender, intersex, queer/questioning, asexual, non-binary, pansexual, grey, two spirit. |
| Iwi | All the iwi that the survivor identifies with. |
| Gangs | Being a patched or prospective gang member or having whänau in a gang. |
| Incarceration | If the survivor has mentioned that they have been incarcerated at any point in their lives. |
| Imputed or imputation | In statistics, imputation is the process of replacing missing data with substituted values.  For our work we used a hot-deck imputation method where the missing value was randomly selected based on matching setting criteria. This is a common imputation method used to replicate values for missing data based on the existing population distribution. |
| Mental  Distress | Someone who positively states that the survivor has been diagnosed with a mental health condition such as personality disorder, bipolar, depression, anxiety, paranoia, obsessive compulsive disorder. Also yes if they have an addiction (alcohol, drugs, gambling) or clearly are abusing substances. This is if this survivor has been diagnosed with any of the above at any point in their life. |
| Chronic health condition | Someone who positively stated that they have a chronic health condition (physical or psychological) such as diabetes, blood pressure, depression, anxiety, heart disease, arthritis. Regardless of if this is directly attributed to their time in care. this is if this survivor has been diagnosed with any of the above at any point in their life. |
| Deaf | Someone who positively states that they are Deaf or have sensory issues affecting their hearing, at any point in their life. |

|  |  |
| --- | --- |
| DisabilityCommunicati on | Someone who positively states that they have a communication and speech impairment at any point in their life such as stutter, impaired articulation, voice impairment or language impairment. |
| Disability Blind | Someone who positively states that they are blind or have sensory issues affecting their sight at any point in their life. |
| Disability  Learning | Someone that identifies as a person with a learning disability at any point in their life. Some people that are neurodivergent may also have learning disabilities (but not always). Learning disability can affect people's:  • Intellectual functioning (such as learning, problem solving, judgement) Adaptive functioning (activities of daily life such as communication and independent living). In transcripts it may be referred to as intellectual disability, intellectually disabled, person with a learning disability, person with learning disabilities. |
| Disability Mobility | Someone that positively states that they have a mobility impairment and/or physical disability such as cerebral palsy, hemiplegia, paraplegia, amputee, or degenerative illness e.g. stroke, Parkinson's disease, at any point in their life. |
| Neurodiverg ent | Broad range of neurological difference(s) or conditions, including Autism, Foetal Alcohol Spectrum Disorders, ADHD, dyslexia, dyscalculia, dyspraxia and dysgraphia. This could be at any point in their life. |
| Physical Abuse | Caregiver hitting child and instances where child has hit other child. |
| Emotional Abuse | Psychological, emotional, mental (include witnessing abuse). |
| Sexual  Abuse | Non-consensual engagement in sexual activities or behaviours, involving physical, emotional, or psychological coercion, manipulation, or force, which causes harm to the victim. |
| Sexual  Non-Contact  Abuse | Non-contact sexual abuse refers to actions that do not involve direct physical contact, yet still involve coercion, manipulation, or force of a sexual nature, For example grooming or exposing someone to explicit content without their consent. |
| Medical  Abuse | Where a survivor has stated overmedication, restraint, seclusion, treatment without consent, sterilisation, forced abortion, conversion therapy, aversion therapy or a misdiagnosis. |
| Neglect | Survivors not having the essential level of physical, emotional or other support, leading to potential harm and compromising the well-being of an individual. |
| Solitary | Where a survivor while in care has been isolated and confined to a small, often windowless cell for a significant portion of the day, with minimal human interaction and restricted sensory stimulation. |

### 10.2. Definitions of Groupings within Variables

For groupings that have been merged or called 'other':

|  |  |
| --- | --- |
| Grouping | Definition |
| Any Abuse | Survivors who have recorded one or more different types of abuse. The types of abuse included under this grouping is - Physical, emotional, sexual, non-contact sexual, medical, neglect and solitary. |
| Gender - Gender Diverse, non-binary or other | Survivor explicitly stated that their gender is "gender diverse, non-binary or other". |
| Ethnicity - Other | Survivor explicitly stated that they identify with an ethnicity other than Päkehä, Mäori, Pasifika, Asianj or MEL AA.  In tables and graphs where no explicit values are shown for Asian and MEL AA, 'Other' group combines the values for Asian, MELAA and Other ethnicities. |

### 10.3. Iwi Classification

The iwi identified by survivors were grouped according to Stats NZ "Iwi and iwi-related groups statistical classification V2.1.0t

|  |  |
| --- | --- |
| Iwi listed by Iwi Groups | Number of Survivors |
| Te Tai Tokerau/Tämaki-makaurau (Northland/AuckIand) Region Iwi | 190 |
| Ngäi Takoto, Ngäpuhi, Ngati Hine, Ngäti Kahu, Ngäti Kahu ki Whangaroa, Ngäti Kurä, Ngäti Kuri, Ngäti Whätua, Ngätiwai, Te Aupäuri, Te Rarawa, Te Roroa |  |
| Confederations and Waka, iwi not named | 127 |
| Tainui, Te Arawa |  |
| Te Tai Räwhiti (East Coast) Region Iwi | 119 |
| Ngäi Tämanuhiri, Ngäti Porou, Rongowhakaata, Te Aitanga ä Mähaki |  |
| Tauranga Moana/Mätaatua (Bay of Plenty) Region Iwi | 114 |
| Ngäi Te Rangi, Ngäti Awa, Ngäti Manawa, Ngäti Pükenga, Ngäti Ranginui, Ngäti Tüwharetoa (Bay of Plenty), Te Whänau a Apanui, Tuhoe, Whakatöhea |  |

|  |  |
| --- | --- |
| Te Matau-a-Mäui/Wairarapa (Hawke's Bay/Wairarapa) Region iwi | 93 |
| Ngäti Hineuru, Ngäti Kahungunu, Ngäti Kahungunu ki Heretaunga Tamatea,  Ngäti Kahungunu ki Te Wairoa, Ngäti Kahungunu ki Wairarapa - Tämaki Nui ä  Rua, Ngäti Pähauwera, Ngäti Rongomaiwahine, Ngäti Ruapani mai  Waikaremoana, Te Wairoa iwi and hapu |  |
| Te Waipounamu (South Island) Region Iwi | 85 |
| Käti Mämoe, Ngäi Tahu, Ngäti Apa ki te Rä Tä, Ngätj Rärua |  |
| Waikato/Te Rohe Pötae (Waikato/King Country) Region Iwi | 70 |
| Ngäti Hikairo, Ngäti Korokä Kahukura, Ngäti Maniapoto, Ngäti Te Wehi, Raukawa, Waikato |  |
| Taranaki Region Iwi | 37 |
| Ngä Rauru Kätahil Ngäruahine, Ngäti Maru (Taranaki), Ngäti Ruanui, Taranaki, Te Atiawa (Taranaki), Te Pakakohi |  |
| Iwi named, region not known | 35 |
| Ngäti Apa, Ngäti Hauä, Ngäti Mutunga, Ngäti Tama, Ngäti Toa Rangatira, Rangitäne, Waitaha |  |
| Manawatü/Horowhenua/Te Whanganui-a-Tara  (Manawatü/Horowhenua/WelIington) Region Iwi | 33 |
| Muaüpoko, Ngäti Raukawa ki te Tonga, Te Atiawa (Wellington) |  |
| Refused to Answer | 28 |
| Don't Know | 29 |
| Whanganui/Rangitikei (Wanganui/RangitTkei) Region Iwi | 20 |
| Ngäti Hauiti, Ngäti Rangi, Te Ati Haunui-a-Päpärangi, Te Korowai o Wainuiärua  (Central Whanganui), Whanganui Iwi / Te Atihaunui a Päpärangi, Whanganui Land Settlement (Lower Whanganui) |  |
| Hauraki (Coromandel) Region Iwi | 19 |
| Ngäti Hako, Ngäti Maru (Hauraki), Ngäti Paoa, Ngäti Porou ki Harataunga ki Mataora, Ngäti Pükenga ki Waiau, Ngäti Tamaterä, Ngäti Whanaunga |  |
| Te Arawa/Taupö (Rotorua/Taupö) Region Iwi | 15 |
| Ngäti Pikiao, Ngäti Rangitihi, Ngätj Rangiwewehif Ngäti Tahu / Ngätj Whaoa, Ngati Whakaue, Tähourangi |  |
| Rékohu/Wharekauri (Chatham Islands) Region Imi/lwi | 6 |
| Moriori, Ngäti Mutunga o Wharekauri (Chatham Islands) |  |

### 10.4. Excluding Accounts Out of Scope

From the CRM extract, the survivors table contained a total of 4,264 accounts. For this analysis, only "survivors with accounts" were requested to be included by the Royal Commision and so only 2,329 survivors have been analysed. The excluded accounts are made up of: 1,336 direct survivors without accounts, 271 impacted survivors including witnesses of abuse, 321 individuals without survivor type and 7 duplicate accounts.

### 10.5. DOT's Data Security and Storing

DOT takes privacy and security seriously, as we understand how paramount it is to ensure the safety and protection of survivor accounts. DOT employed a number of security measures to ensure this throughout this project and applied security principles to ensure privacy and security when data is in transition, at rest or actively used within DOT I s infrastructure.

Data security measures were implemented during the transition period to ensure the protection of sensitive information. Encrypted connections were used for the data flow between the Royal Commission and DOT, and a VPN was employed to enable remote access to the internal network for staff working remotely. Early Warning Systems were in place to actively scan ports, check web certificates, and identify breached email addresses. Additionally, a Meraki Firewall implemented white-listing and extensive logging to restrict access to authorised clients for on-premises services. Security Groups acted as a firewall within AWSI controlling inbound and outbound traffic for associated instances.

During active data usage, strict protocols were followed to maintain data security. Data was only accessible to designated principal and secondary investigators and was used solely for analysis purposes while respecting privacy and confidentiality. Microsoft Active Directory and Security Groups systems ensured controlled access to Windows-based file servers. Database servers relied on a role-based security model, and isolated docker environments were employed for each data store to minimise the risk of compromising multiple databases in the event of unauthorised access.

For data at rest, Advanced Encryption Standard (AES-256) encryption was applied at the point of ingestion. Secure wiping processes were implemented to remove client data after completion of specific tasks, and equipment decommissioning involved thorough data wiping to eliminate any residual traces. Throughout the data processing phase, anonymization techniques were utilised to further protect sensitive information.

#### 10.6. Methodology

DOT's project methodology drew upon our expertise in natural language processing projects and comprised five key stages:

##### 10.6.1. Data Loading and Preprocessing

We assessed data quality by analysing the distribution of missing information in Arahiko CRM data. Audio and non-textual files (e.g., .msg files) were excluded from further processing. We prototyped methods to handle missing information and estimated project costs and timelines. Regular meetings with the Royal Commission helped determine data field scope and priority, categorised as required or optional.

##### 10.6.2. Text Processing and Tokenization

Individual documents were loaded from their source files and combined into a single dataset. Each word in the text was assigned a part-of-speech (e.g., noun, verb, adjective). Sentences were split based on part-of-speech tags, differentiating between self-reference and others' references. Sentences were used to populate the fields in our criteria.

##### 10.6.3. Keyword Tagging

Collaborating with the Royal Commission, we created an appropriate synonym list. Sentences were categorised based on the presence of identified keywords.

##### 10.6.4. Data Modelling

We employed the language model BERT to classify sentences into positive and negative observations, considering context and meaning. Models were trained for each field using manually labelled data from stage (3). Model performance was tested using three accuracy criteria:

* Accuracy: A measure of overall model correctness, calculated as the percentage of correct predictions out of all instances.
* Precision: The accuracy of positive predictions made by the model.
* Recall: Measures model sensitivity, calculating the proportion of correctly predicted positive instances out of all positive instances.

Priority was given to models with high precision, and poorly performing models were adjusted to reduce inaccuracies. The final models were applied to all tagged outputs from stage (3).

##### 10.6.5. Prediction Review

We reviewed the model predictions, making necessary corrections. Depending on the required granularity, evidence was either grouped by RC number or reviewed on a sentence level. High numbers of false positives associated with the same keyword led to updates of synonym lists or model retraining.

DOT I s overall methodology involved a systematic review of data quality followed by the application of NLP techniques to enhance the data. The approach included collaboration with the Royal Commission, regular feedback loops, and the use of deterministic and probabilistic methods for data extraction and deduplication. Transparency, auditability, and accuracy were emphasised throughout the process, and the resulting data was structured and suitable for subsequent analyses.

DOT then manually reviewed 100% of 26 out of 29 attributes and undertook a random sample of between 25-50% for the remaining four attributes. These were: Physical Abuse, Sexual Abuse, Emotional Abuse and Institution.

DOT informed AIC that, due to very tight timeframes, DOT was not able to manually check 100% of the positive cases for these four attributes. The agreed standard with AIC was a precision of 90% where precision is the true positive rate. This was achieved in the samples based on independent review by AIC.

##### 10.6.6. Data Imputation

In the datasets containing information about abusive events and time spent in care, we encountered significant proportions of missing data. Specifically, in the time in care data, approximately 50% of survivors had missing start and finish age information from their time in an institution. Additionally, among the survivors for whom we had information on abusive events, 19% were missing start and end age details for those events. Finally, in 20% of cases, the decade during which the abuse occurred was not identified.

To address these gaps, we employed a hot deck imputation methodology. This approach entails selecting ages for survivors within a given setting by drawing from the age distribution of other survivors within the same setting. Subsequently, we applied a similar methodology to impute ages during which abuse took place. These imputation processes enabled us to estimate missing data points.

Coupling our imputed age data with respective survivor birth years, we calculated an estimation for the decade in which the abuse occurred.

#### 10.7. Supplementary Tables

10.7.1. Survivors who have been incarcerated by Gang Associations

|  |  |  |
| --- | --- | --- |
| Gang Association | Survivors who have been incarcerated | Survivors who have never been incarcerated |
| Gang Member | 145 | 84 |
| Gang Whänau | 50 | 54 |
| No Association | 488 | 1508 |
| Total | 683 | 1646 |

10.7.2. Survivor Count by Pathway into Care

|  |  |  |
| --- | --- | --- |
| Pathway into Care | Count | Percentage |
| Taken away by CYFS | 172 |  |
| Parents' voluntary placement | 91 |  |
| Faith-based school | 33 |  |
| Psychiatric | 28 |  |
| Not in full-time care | 20 | 0.9% |
| Tragedy | 15 | 0.6% |
| Disability | 8 | 0.3% |
| Unwed pregnancy/Faith community |  | 0.3% |
| Unidentified | 1955 | 84% |
| Total (accounted for) | 374 |  |