Australian Government

Australian Institute of Criminology

# CRIME & JUSTICE RESEARCH 2026

Edited by Rick Brown

Australian Institute of Criminology

# CRIME & JUSTICE RESEARCH 2026

Edited by Rick Brown

© Australian Institute of Criminology 2026

Apart from any fair dealing for the purpose of private study, research, criticism or review, as permitted under the *Copyright Act 1968* (Cth), no part of this publication may in any form or by any means (electronic, mechanical, microcopying, photocopying, recording or otherwise) be reproduced, stored in a retrieval system or transmitted without prior written permission. Inquiries should be addressed to the publisher.

Published by the Australian Institute of Criminology
GPO Box 1936
Canberra ACT 2601
Tel: (02) 6268 7166
Email: **front.desk@aic.gov.au**

Edited by Rick Brown, Deputy Director, Australian Institute of Criminology

ISBN 978 1 922878 37 3 (Print)
ISBN 978 1 922878 38 0 (Online)
**https://doi.org/10.52922/sp78380**

For a complete list of AIC publications, visit the Australian Institute of Criminology website: **www.aic.gov.au**

Disclaimer: This research does not necessarily reflect the policy position of the Australian Government.

# Contents

**About the editor** **vi**

**List of contributors** **vii**

**Introduction** **1**

**1. Validity of the Violence Risk Scale for male prisoners in Australia** **4**
*Emma Ziersch, Shawn Sowerbutts, Yilma Woldgabreal and Sophie Ransom*

**2. Supporting incarcerated mothers: A mixed methods evaluation of the NSW Co-Located Caseworker Program** **19**
*Althea Gibson, Marc Rémond, Peta MacGillivray, Eileen Baldry and Elizabeth Sullivan*

**3. The extent of illicit drug and alcohol involvement in crime: An updated estimate** **34**
*Alexandra Voce, Tom Sullivan, Anthony Morgan and Merran McAlister*

**4. An experimental study of support for protest causes and tactics and the influence of conspiratorial beliefs** **48**
*Anthony Morgan, Timothy Cubitt, Alexandra Voce and Isabella Voce*

**5. Exposure to and sharing of fringe or radical content online** **68**
*Timothy Cubitt and Anthony Morgan*

**6. The overlap between viewing child sexual abuse material and fringe or radical content online** **84**
*Timothy Cubitt, Anthony Morgan and Rick Brown*

**7. Artificial intelligence and child sexual abuse: A rapid evidence assessment** **98**
*Heather Wolbers, Timothy Cubitt and Michael John Cahill*

**8. Drivers and deterrents of child sexual offending: Analysis of offender interactions on the darknet** **113**
*Heather Wolbers, Timothy Cubitt, Michael John Cahill, Matthew Ball, John Hancock, Sarah Napier and Roderic Broadhurst*

**9. Financial risk indicators of child sexual abuse live streaming: A proof of concept prediction model** **127**
*Timothy Cubitt, Sarah Napier and Rick Brown*

**10. Sexual extortion of Australian adolescents: Results from a national survey** **142**
*Heather Wolbers, Timothy Cubitt, Sarah Napier, Michael John Cahill, Mariesa Nicholas, Melanie Burton and Katherine Giunta*

**11. The impacts of sexual extortion on minors: A systematic review** **164**
*Heather Wolbers, Timothy Cubitt, Renee Carter and Sarah Napier*

**12. Underage use of adult-based platforms: Risks and experiences** **178**
*Tahlia Hart*

**13. Victim-survivors' reflections on best practice in restorative justice for domestic, family and sexual violence** **191**
*Siobhan Lawler*

**14. Ransomware targeting individuals and small businesses: Vulnerabilities and impacts** **204**
*Isabella Voce and Anthony Morgan*

**15. Developing a harm index for individual victims of cybercrime** **220**
*Isabella Voce and Anthony Morgan*

**Index** **238**

# About the editor

Dr Rick Brown is Deputy Director of the Australian Institute of Criminology, where he oversees the day-to-day work of the Institute.

Rick has been a criminologist for over 35 years and has worked extensively on issues associated with crime prevention, community safety and policing in Australia, the United Kingdom and Republic of Ireland. He was previously the Managing Director of a research consultancy that focused on crime and justice issues and worked as a Research Manager for the Home Office in London, where he ran programs of research on vehicle crime and domestic burglary.

Rick has published over 80 government reports, book chapters and journal articles on a range of topics including antisocial behaviour, arson, burglary, CCTV, community justice, crime prevention, domestic violence, drink-driving, drug use, identity crime, organised crime, policing and vehicle crime. Most recently he has worked on issues associated with child sexual abuse material, publishing a book with Routledge in 2022 titled *Eliminating online child sexual abuse material*.

He holds a PhD from the London School of Economics and Political Science, a Master of Arts degree from the University of Westminster and a Bachelor of Arts (with Honours) degree from Hatfield Polytechnic in the United Kingdom.

# List of contributors

**Eileen Baldry** AO, FASSA, FRSNSW, FANZOC (PhD) is Professor Emerita of Criminology, UNSW. Her research focuses on social justice including criminalisation of mental and cognitive disability; criminalised women, girls and First Nations people; homelessness and social housing. She is a chief investigator on Australian Research Council, National Health and Medical Research Council and Australian Housing and Urban Research Institute grants, contributes to commissions and inquiries and has published over 170 peer-reviewed books, articles and reports.

**Matthew Ball** is the Laboratory Coordinator of the Australian National University's Cybercrime Observatory. His research explores the underworld of the internet. In particular, his work investigating online darknets aims to reveal information on the hidden populations and subcultures operating online.

**Roderic Broadhurst** PhD was an Emeritus Professor and Fellow in the Research School of Asia and the Pacific at the Australian National University (ANU). He was Director of the ANU Cybercrime Observatory and a former Professor of Criminology and Deputy Director the Australian Research Council Centre of Excellence in Policing and Security. He was foundation editor of the *Asian Journal of Criminology* and wrote widely on topics such as crime statistics, recidivism, crime victims, organised crime and cybercrime.

**Melanie Burton** is a Senior Research Officer at the eSafety Commission with over a decade of experience in mixed-methods research focused on online safety. Her expertise includes technology-facilitated abuse, cyberbullying, image-based abuse, and child sexual exploitation. Her work supports evidence-informed policy, programs and regulatory functions, strengthening understanding of evolving online harms. Previously, she researched family, domestic and sexual violence. Melanie has extensive experience engaging diverse and vulnerable populations, is a former child and adolescent clinical psychologist, and holds a PhD in Law (Criminology) from UNSW.

**Michael John Cahill** is a former Research Analyst at the Australian Institute of Criminology.

**Renee Carter** is a former intern at the Australian Institute of Criminology.

**Timothy Cubitt** is an Associate Professor of Criminology at Adelaide University and Deputy Director of the Australian Cybercrime Observatory. Tim's research focuses on child sexual exploitation and abuse, extremism, and policing. In particular, he examines the dynamics of offending and victimisation, and offender behaviour and organisation, with a focus on technological and financial enablers. Prior to joining Adelaide University, he worked with the New South Wales Police Force and the Australian Institute of Criminology.

**Dr Althea Gibson** is a Lecturer in the Faculty of Law and Justice, UNSW Sydney. Her areas of research include domestic and family violence law and policy, evidence law and the intersection of criminal justice and child protection systems. Her recent research focuses on the criminalisation of women's resistance to domestic and family violence and approaches to risk assessment and management within the domestic and family violence sector.

**Katherine Giunta** is a qualitative researcher specialising in research with young people and queer communities. With a background in youth mental health, gender studies and anthropology, she designs and delivers trauma-informed research centred on lived experience. She has extensive experience working on sensitive topics, including mental health and pornography. Katherine has led the development of distress and disclosure protocols at the eSafety Commission and holds a PhD in Anthropology, alongside relevant youth and suicide intervention qualifications.

**John Hancock** is a former Research Assistant at the Australian National University's Cybercrime Observatory. He has conducted research into the processes of alt-right radicalisation and the behaviours of offenders in child sexual abuse material forums online.

**Dr Tahlia Hart** is a practising lawyer and a lecturer in Criminology at Flinders University's College of Business, Creative Arts, Law and Social Sciences. Her research primarily focuses on children's learnt skills and experiences in online environments intended for adult users, such as dating apps. Tahlia's broader research interests include technology, cybercrime, and digital justice for vulnerable cohorts to better understand and mitigate potential risks and harms.

**Siobhan Lawler** is a Principal Research Analyst at the Australian Institute of Criminology, where she leads programs of research on elder abuse and sexual violence. Siobhan holds a PhD in Public Health from the University of Sydney and a Bachelor of Psychological Science and Criminology (Honours) from the University of New South Wales.

**Peta MacGillivray** (Kalkutungu and South Sea Islander) is a Senior Research Fellow at UNSW Sydney in the Social Determinants of Justice Research Hub, whose work focuses on transforming public policy through Indigenous-led partnerships and community-based research. Recent research focuses on evidence-informed pathways to improve social outcomes for children and young people impacted by government systems.

**Merran McAlister** is a Senior Research Analyst in the Australian Institute of Criminology's Serious and Organised Crime, Cybercrime and Radicalisation Research Program. She obtained a Bachelor of Criminology, Bachelor of Science (Psychology) and Master of Criminology, Justice and Regulation from the Australian National University. She has published in the areas of fraud and corruption, sexual offending, deaths in custody, and identity crime and misuse.

**Anthony Morgan** is the Research Manager of the Australian Institute of Criminology's Serious and Organised Crime, Cybercrime and Radicalisation Research Program. Committed to promoting evidence-based policy, Anthony has spent two decades working closely with law enforcement and policymakers to conduct rigorous, applied and policy-relevant empirical criminological research.

**Dr Sarah Napier** is Research Manager of the Online Sexual Exploitation of Children Research Program at the Australian Institute of Criminology. The purpose of Sarah's research is to understand offending behaviour and the needs of survivors to inform prevention and disruption of child sexual exploitation and abuse. Sarah holds a PhD in Criminology from the University of Sydney, and her doctoral research focused on pathways to child sexual abuse material offending.

**Mariesa Nicholas** is an experienced research leader overseeing the research and evaluation function at the eSafety Commission. She specialises in lived experience research on sensitive and complex issues, including online harms, mental health and help-seeking. Drawing on a strong background in qualitative and participatory research, she has developed methodologies that support participant comfort and safety when exploring stigmatised topics. Mariesa holds a double BA in Social Science and International Studies.

**Sophie Ransom** has extensive experience in research and evaluation across the justice, public health and community service sectors. Her research interests include therapeutic responses to justice-involved people, domestic violence reduction and evidence-based policy development. Sophie works at the South Australian Department for Correctional Services, where her previous work includes evaluations of rehabilitation, cultural and employment programs for correctional clients and validation studies of criminogenic risk assessment tools.

**Dr Marc Rémond** (Senior Research Officer, University of Newcastle) has a background in health research, law and biology. Most recently, Marc's work has focused on justice health research, including projects relating to intergenerational incarceration, throughcare, deaths in custody, and mothers in custody. He also collaborates on research projects relating to Aboriginal and Torres Strait Islander health (particularly rheumatic heart disease) and rare and serious conditions in pregnancy (notably cardiac diseases and cancer).

**Shawn Sowerbutts** currently works as a Senior Psychologist within the NDIS Team, Department for Correctional Services in South Australia, where he facilitates access to NDIS services for eligible clients. Prior to this, he worked as the Senior Programs Officer in the Program Services Unit, a role in which he developed criminogenic treatment programs. A member of the department's Reconciliation Action Plan Working Party, Shawn is committed to providing more culturally appropriate services to Aboriginal people.

**Elizabeth Sullivan** is a public health physician, executive leader and researcher. She is Professor of Equity and Health at the University of Newcastle and Research Lead Custodial Health at the Justice Health and Forensic Mental Health Network. She leads a program of health and social services research and evaluation focused on addressing the impact of the criminalisation of poverty on mothering and children and the intergenerational transmission of incarceration.

**Tom Sullivan** is a Principal Research Analyst in the Australian Institute of Criminology's Statistical Collections and Indigenous Justice Research Program. Tom has led research into sexual offending, illicit drugs and crime prevention. He has experience in qualitative methods, the design and management of surveys and the collection and analysis of administrative data. He holds a Bachelor of Behavioural Science and Master of Arts from Flinders University.

**Dr Alexandra Voce** is an Acting Principal Research Analyst in the Australian Institute of Criminology's Serious and Organised Crime, Cybercrime and Radicalisation Research Program. She leads the program of work on transnational serious and organised crime and publishes a yearly report estimating the cost of serious and organised crime to Australia. Alexandra has a PhD in Psychology, with her doctoral research focusing on the link between illicit drug use and mental health.

**Isabella Voce** is a Principal Research Analyst for the Australian Institute of Criminology's Serious and Organised Crime, Cybercrime and Radicalisation Research Program. Isabella works closely with law enforcement and policymakers to conduct rigorous, empirical criminological research to better understand and disrupt cybercrime. This includes understanding national cybercrime trends, and what works in cybercrime prevention and victim support.

**Dr Heather Wolbers** is a Principal Research Analyst in the Family, Domestic and Sexual Violence, and Human Trafficking and Modern Slavery Research Programs at the Australian Institute of Criminology. Heather has undertaken research in the fields of family violence, sexual violence, the sexual exploitation of children, and sexual extortion of young people. Heather holds a Bachelor of Criminology and Criminal Justice (with Honours), a Bachelor of Forensic Science, and a PhD in Criminology from Griffith University.

**Yilma Woldgabreal** has a PhD in Psychology and more than 25 years of experience with South Australia's Department for Correctional Services. In 2024, he was honoured with the National Australian Corrections Medal in recognition of his dedicated service across roles from frontline positions to senior management. He is an Adjunct Academic at Flinders University and has published broadly on risk assessment, rehabilitation, reintegration initiatives, ethnic minorities in the justice system, and strength-based intervention strategies.

**Emma Ziersch** is a Senior Programs Evaluation Officer at the South Australian Department for Correctional Services. During the past eight years Emma has undertaken research and evaluation on a range of programs and initiatives related to offender rehabilitation. Projects include various bail, housing and employment support initiatives, and reoffending studies of the department's general violence and domestic violence programs.

# Introduction

This volume provides a snapshot of recent research published by the Australian Institute of Criminology (AIC). Each year, the AIC produces around 50 publications. These include peer-reviewed Research Reports and *Trends & issues in crime and criminal justice* papers, as well as Statistical Reports, Statistical Bulletins and Research in Practice papers—all of which provide insight into contemporary issues faced by the community and the criminal justice system in Australia. Fifteen studies are included in this volume, covering issues as varied as Indigenous justice; transnational serious and organised crime; violent extremism; online sexual exploitation of children; family, domestic and sexual violence; and cybercrime. These themes reflect the AIC's focus on reducing crime and promoting justice by improving evidence-based policymaking, with research examining some of the most pressing crime and justice issues affecting Australia.

Chapter 1 presents the findings from a study examining the validity of the Violence Risk Scale—a statistical tool typically used in correctional settings to assess the risk of violence among newly arrived prisoners. Drawing on a sample of 2,689 assessments from multiple jurisdictions in Australia, the study examined the risk profile of male prisoners (both Aboriginal and non-Indigenous) according to the Violence Risk Scale and compared these with reoffending outcomes. The results showed that the scale was much more likely to rate Aboriginal men as being at high risk of reoffending compared with non-Indigenous men. Worryingly, the scale also produced a high rate of false positives for Aboriginal men—many were classified as being at high risk of reoffending when that was not in fact the case. This could lead to them being imprisoned under stricter conditions and being less likely to receive parole. This demonstrates how a seemingly objective tool could lead to adverse outcomes for Aboriginal men.

In contrast, chapter 2 examines the positive outcomes experienced by those participating in the New South Wales Co-Located Caseworker Program, which was established to support women in custody who have children involved in the child protection system. By co-locating child protection caseworkers in correctional centres, the program improved outcomes for both mothers and their children by providing a professional who could liaise with outside services and advocate for the family on child protection matters. There was general agreement on the merits of the program among the research participants, who recognised the benefits that could accrue from rolling out the initiative to other correctional centres.

Produced under the AIC's transnational, serious and organised crime research program, the study presented in chapter 3 analyses the relationship between drug and alcohol use and engagement in crime. Based on interviews with more than 2,200 police detainees, undertaken as part of the AIC's former Drug Use Monitoring in Australia program, the research examines the extent to which those who use alcohol and other drugs also engage in crime attributed to that substance use. It finds that substance use contributed to more than 156,000 principal offences in 2019. Almost 96,000 of these offences were attributed to methamphetamine use alone. This is concerning in the context of increasing methamphetamine use in Australia. Indeed, the proportion of police detainees who attributed their offending to methamphetamine increased from 6.2 percent in 2009 to 27.6 percent in 2019.

The AIC has undertaken a program of research on violent extremism in recent years, examining factors that increase the propensity for radicalisation. Chapter 4 reports on the results of a randomised survey experiment in which respondents were asked about their support for environmental, anti-lockdown and sovereign citizen protests. They were randomly allocated to one of three groups presented with different protest tactics—peaceful marching, disrupting traffic and violent clashes with police. Encouragingly, the research found that respondents were significantly more likely to oppose violent or disruptive protests than peaceful protests, regardless of the issue or movement in question. However, the effect of conspiratorial beliefs on support for protests was found to vary by protest cause. Belief in conspiracy theories increased support for protest violence, relative to other tactics, and support for certain protest causes and tactics was shaped by a person's ideological beliefs.

Chapter 5 explores the extent to which Australians access and share fringe and radical content online. In an online survey of over 13,000 Australian adults, 40.6 percent reported being exposed to material they described as fringe, unorthodox or radical. A quarter of these respondents accessed the content intentionally and a third reported that the content depicted violence. Fringe and radical content was most frequently shared on mainstream social media and messaging platforms.

Exploring the intersection between two AIC research programs (those on violent extremism and online sexual exploitation of children), chapter 6 reports on a study measuring the extent to which those accessing fringe or radical content online also viewed child sexual abuse material (CSAM), finding a significant association between the two. Among survey respondents who viewed CSAM, 64.7 percent had also viewed fringe or radical content, while 7.1 percent of those who viewed radical content had also viewed CSAM. Individuals who viewed both kinds of material online were more likely than other respondents to be young and male and to have had criminal justice system contact and diagnosed mental illness.

The AIC has also recently produced numerous reports examining other aspects of online sexual exploitation of children. For example, chapter 7 reviews the intersection of artificial intelligence (AI) and child sexual abuse, in recognition of the growing trend towards AI-generated CSAM over the past few years. Although still an emerging field of inquiry, 33 empirical studies were identified. However, all of these studies explored ways AI can assist with law enforcement investigation and detection, and no studies examined how AI was used to generate CSAM, likely due to the recency of the problem.

Chapter 8 reports the findings of a study that examined a darknet forum for undetected online and contact child sexual offenders to identify key drivers and deterrents of offending and to inform intervention approaches. It found that forum participants normalised sexual contact with children while minimising or denying the resulting harm and shifting the responsibility for offending. Conversely, acknowledgement of the harm to children, feelings of guilt and shame, and concern about being caught by law enforcement or detected by family and friends acted as deterrents to continued offending.

Chapter 9 reports on an innovative use of machine learning, applying this tool to financial transaction data to identify Australian perpetrators sending money to the Philippines to purchase the live streamed abuse of children. The research showed that, with a very limited set of financial indicators, about half of the live streaming perpetrators in Australia could potentially be identified. Importantly, this could be achieved with a very low rate of false positives, with just one in 300 cases involving a non-offender being misclassified as an offender. This model could potentially be used by financial institutions to identify and block transactions being made for the abuse of children with a greater degree of confidence than is currently the case.

Chapter 10 examines the victimology associated with sexual extortion of young people. Based on an online survey of 16- to 18-year-olds, the research examined the prevalence of victimisation, finding that about one in 10 adolescents had experienced sexual extortion in their lifetime, and one in three victims had experienced it more than once. More than half of victims had experienced sexual extortion before the age of 16, and two in five were extorted using digitally manipulated material. Two-thirds of those who experienced sexual extortion had only ever met the perpetrator online. Chapter 11 also examines sexual extortion of young people, although focusing on the impacts experienced, including psychological harm, socio-environmental impacts, financial costs, risk of ongoing victimisation, and criminal justice involvement. Taken together, chapters 10 and 11 demonstrate both the extent of the problem and the enduring harm sexual extortion can cause to those affected.

Chapter 12 explores the risks that young people take when using platforms intended for adults (such as dating apps, remunerative matchmaking services, subscription platforms, fetish communities). Based on interviews with 46 young adults about their risk-taking experiences as children on these platforms, the study documents a range of ways in which children are exposed to risk both online and offline. This included inappropriate or illegal interactions involving sexting and activities such as drug use.

Chapter 13 reports on a study from the AIC's domestic, family and sexual violence research program. This study focuses on the use of restorative justice in domestic and family violence and sexual violence cases, exploring best practice principles from the perspectives of victim-survivors. Key findings included the importance of Convenors' skills and expertise, flexibility and a feeling that victim-survivors were supported to satisfy a diverse range of justice needs. These were considered critical to the success of restorative justice in this context.

Two reports from the AIC's cybercrime research program have also been included. The first, presented in chapter 14, examines vulnerabilities to and impacts of ransomware victimisation among small to medium enterprise owners. It shows that victims had often received multiple ransom messages in the past 12 months, and many had previously paid a ransom. The second study, presented in chapter 15, creates a detailed cybercrime harm index that can be used to compare the impacts on the Australian community associated with 17 different types of cybercrime. Based on 34 dimensions of harm—encompassing practical, health, social, financial and legal impacts—the study identifies stalking and harassment as the most harmful form of cybercrime in Australia, followed by remote access scams and stealing or sharing content without consent. Importantly, the research also demonstrates how cybercrime harms are concentrated among certain segments of society, with just 10.9 percent of victims experiencing 57.7 percent of all harm.

This volume contains a small fraction of the research the AIC has undertaken in recent times. Each study presented here provides an empirical foundation for policymaking. Given the ever-changing nature of crime, these topics are constantly evolving to ensure the AIC's work continues to be contemporary and relevant. Further details of the AIC's research can be found at **aic.gov.au**.

**Dr Rick Brown**

**Deputy Director**
**Australian Institute of Criminology**

# 1. Validity of the Violence Risk Scale for male prisoners in Australia

Emma Ziersch, Shawn Sowerbutts, Yilma Woldgabreal and Sophie Ransom

The over-representation of Aboriginal and Torres Strait Islander Australians in the criminal justice system as a result of colonisation is well-documented and remains a critical issue despite the introduction of various strategies to address the problem. For example, recent data from the Productivity Commission (2022) on the progress of Target 10 under the National Agreement on Closing the Gap revealed that from 2019 to 2021 the over-representation of Indigenous adults in the criminal justice system worsened. Figures from the Australian Bureau of Statistics (ABS) showed that First Nations people represented 3.8 percent of the overall Australian population yet comprised 31.8 percent of adults in custody. Moreover, 77.7 percent of Indigenous people in custody had experienced prior adult imprisonment (ABS 2022, 2021).

For individuals sentenced to correctional supervision, risk assessment tools form the basis for decision-making around how they progress through the correctional system. These tools are used to determine likelihood of further offending behaviour and support the ability of treatment providers to place individuals in treatment of the correct intensity and that addresses relevant criminogenic needs. For individuals convicted of offences involving violence, most jurisdictions in Australia use the Violence Risk Scale (VRS; Wong & Gordon 2003). This tool is based on the risk-need-responsivity (RNR) model (Andrews, Bonta & Hoge 1990) and is designed to assess and predict risk of future offending, inform decisions around therapeutic intervention and measure changes in violence risk as a result of treatment (Wong & Gordon 2006).

However, as most instruments for estimating risk of reoffending have been normed on samples largely comprising Caucasian participants, their application to diverse populations who may experience unique, culturally specific risk factors has been criticised (Shepherd et al. 2014). Critics argue that the assumption that these instruments are generalisable to all is ethnocentric and contributes to the misclassification of minority groups, resulting in further disadvantage (Mals et al. 2000; Martel, Brassard & Jaccoud 2011; Westerman & Dear 2023; Woldgabreal, Day & Tamatea 2020). Indeed, some researchers contend that many of the central eight risk markers in the RNR model fail to consider the significant impact of colonisation and intergenerational trauma that influences Indigenous offending trajectories (Allan & Dawson 2004; Jones et al. 2002; Shepherd & Lewis-Fernandez 2016).

Woldgabreal, Day and Tamatea (2020) argue that the theoretical underpinnings of contemporary risk assessments have an individualistic lens and largely ignore the broader historical and structural drivers of crime. They assert that this further exacerbates the over-representation of people of colour given these tools are based on criminological theories that emphasise personal responsibility and fail to consider any inherent cultural bias. These tools are used to inform decisions around sentencing pathways in corrections (eg decisions around suitability for community-based supervision, prison security ratings, and suitability for conditional release), which can further disadvantage First Nations people (Woldgabreal, Day & Tamatea 2020). Areas of concern they identified included misclassification, over-assessment, placement in the wrong rehabilitation pathways, and longer than necessary imprisonment and/or supervision.

Related to this is the dilemma that in applying a uniform, objective risk instrument to all offending populations (under the assumption that risk markers do not vary as a function of ethnicity), the inherent biases within the tools may serve to inflate an Indigenous person's risk status (Martel, Brassard & Jaccoud 2011). That is, in the process of applying a neutral risk assessment to Aboriginal offenders, under contemporary risk assessment theories' assertions that risk markers are the same across populations, these tools can result in the classification of Aboriginal people in the high-risk, high-need category more often than non-Indigenous populations (Martel, Brassard & Jaccoud 2011). Evidence exists that not only is the accuracy of risk instruments lower for Indigenous populations than for white Anglo populations, but Indigenous individuals regularly receive higher scores on these instruments than non-Indigenous offenders (Shepherd 2018).

There is some support for the argument of inflated risk, with a review of the Australian literature on risk assessments revealing the presence of higher baseline scores for Indigenous participants compared with non-Indigenous participants on tools such as the Level of Service Inventory-Revised (LSI-R; Shepherd et al. 2014). For example, Aboriginal offenders in New South Wales scored consistently higher on every subscale of the LSI-R compared with non-Indigenous offenders and recorded significantly greater criminogenic needs related to education and employment, family discord, constructive leisure time and antisocial peers (Hsu, Caputi & Byrne 2010). A later study on the same population found similar results, with Aboriginal offenders recording higher scores on all subscales of the LSI-R excluding the emotional/personal scale (Watkins 2011). Despite these differences, the tool was found to have similar predictive abilities for both groups.

In relation to violence risk assessment more specifically, a recent study using South Australian correctional data from 366 VRS assessments showed Aboriginal males scored significantly higher than non-Aboriginal males at pre-treatment and post-treatment (Woldgabreal et al. 2022). In addition, the study noted a significantly elevated score for incarcerated Aboriginal men on most of the VRS risk domains. This was consistent with the findings from a prison sample in Western Australia involving LSI-R and VRS Screener Version (VRS-SV) assessments (Gower, Morgan & Saunders 2022). Aboriginal offenders (male and female) who committed violent offences were predominantly identified as high risk on the VRS-SV and, on average, scored higher than their counterparts on all components of this screening tool. The authors highlighted parallels with other Australian research (Shepherd et al. 2014) in relation to Aboriginal offenders having significantly higher scores in the areas of substance abuse and education and employment, factors that have been identified as symptoms of colonisation and intergenerational trauma.

These Australian studies point to cultural disparities in overall risk levels and markers of treatment need. What is not yet clear from current research is whether, and to what extent, these differences impact the ability of the VRS to predict future violence among Australia's First Nations peoples.

## Aim and method

This study examines the predictive and discriminative validity of the VRS for Australian males convicted of a violent offence, in particular Aboriginal individuals. The term Aboriginal is used in describing the current study as the sample is almost entirely based on this population. The use of the term Aboriginal in this context is intended to be inclusive of the very small number of Torres Strait Islander people in the sample. The term 'non-Indigenous people' is used for people who are not the First Peoples of the lands on which they reside. In Australia, this refers to people who are not Aboriginal, Torres Strait Islander, or Aboriginal and Torres Strait Islander.

## Sampling and data collection

Correctional agencies in five jurisdictions—New South Wales, Victoria, Western Australia, South Australia and Tasmania—contributed retrospective VRS item data. The dataset was comprised of males imprisoned for a violent offence who were assessed for violence risk using the VRS tool. Data on the reoffending outcomes of those assessed were supplied by correctional agencies and the New South Wales Bureau of Crime Statistics and Research.

## Description of the Violence Risk Scale

The VRS consists of six static and 20 dynamic predictors. The static variables are current age, age of first violent conviction, number of juvenile convictions, violence throughout life span, prior release failures/escapes, and stability of family upbringing. These should predict reoffending but are largely historical and cannot be influenced by interventions. The dynamic variables include violent lifestyle, criminal personality, criminal attitude, work ethic, criminal peers, interpersonal aggression, emotional regulation/control, violence during institutionalization, weapon use, insight into violence, mental illness, substance abuse, stability of relationships, community support, released to high risk situations, violence cycle, impulsivity, cognitive distortion, compliance with supervision, and security level of release institution. These dynamic variables are considered changeable through factors such as treatment intervention, and any shifts in these domains are linked to changes in violence risk.

All 26 variables are rated on a four-point scale (from 0 to 3) based on case file review and a semi-structured interview, with higher scores on each variable indicating greater links to violence in everyday functioning. The total VRS score provides an assessment of the level of violence risk, with higher scores indicating greater risk and greater priority for intervention.

## Definitions

Recidivism was defined as any conviction for a new violent offence following the index release from prison that resulted in a sanction administered by correctional services. Recidivism rates were examined over equal follow-up periods of two and five years. Individuals with insufficient follow-up time were excluded from the recidivism analyses.

## Sample

The sample consisted of 2,689 adult males incarcerated for a violent offence who were administered a full VRS assessment pre-treatment between 2010 and 2022. Participants were aged between 18 and 82 years at the time of assessment, with the average age similar across jurisdictions ($M$=33 years, $SD$=9.0). The combined VRS sample consisted of 817 Aboriginal males (28.4%), which was representative of the general prison population in Australia (30.3% Aboriginal; ABS 2021). Fifteen individuals in the overall sample (0.5%) identified as Torres Strait Islander or both Aboriginal and Torres Strait Islander. Predictive and discriminative validity analyses were based on a sample of 2,125 males with a two-year follow-up and 938 males with a five-year follow-up.

## Measuring predictive validity

To examine the predictive validity of the tool, discrimination (relative predictive accuracy) and calibration (absolute predictive accuracy) statistics were used. To assess discrimination, receiver operating characteristic (ROC) analyses were used to generate the area under the curve (AUC), a value ranging from 0 to 1 representing the probability that a randomly selected recidivist will have a higher score than a randomly selected non-recidivist. As a general rule, AUC values of 0.56 to less than 0.64 are interpreted as low, 0.64 to less than 0.71 as medium, and 0.71 or greater as high (Rice & Harris 2005). Cohen's *d* was computed to provide a further estimate of the effect size difference between recidivists and non-recidivists on their VRS score. The analysis followed Cohen's (1988) guideline, in which a *d* value of 0.20 was considered small, 0.50 was considered medium and 0.80 was large. A series of logistic regressions were also conducted to examine the relationship of VRS scores to violent recidivism by Aboriginality.

Additional measures of discriminative accuracy were undertaken. These included classification accuracy (ie the total proportion of correct risk ratings); sensitivity (ie the proportion of violent recidivists identified as high risk); and specificity (ie the proportion of non-recidivists identified as low or moderate risk). Although risk is best understood on a continuum, in reality correctional agencies use thresholds to inform many operational decisions. This analysis is based on the dichotomy of high risk versus not high risk (low or moderate risk) and recidivism rates for these two groups. To assess calibration, the expected/observed (E/O) index was calculated to compare recidivism rates with those from the sample used by Olver et al. (2022).

## Limitations

The dataset was limited to individuals who had served a period of imprisonment for a violent offence and had a minimum sentence length (and seriousness of offence) that met the criteria for a full VRS assessment. The use of a custodial sample also meant that it was not possible to make comparisons with a community sample, which might have increased the range in the risk scores.

The sample was generated in a field setting and it was not possible to establish the inter-rater reliability of the sample. Relatedly, there may also be errors in data input, along with issues related to the extraction of VRS, offence and treatment information from multiple states. While efforts were made to minimise inaccuracies through the development of a data collection protocol and agreed definitions of terms, the use of data from different agencies may have resulted in some inconsistencies.

A further limitation of the study was the exclusion of female violent offenders due to not all states administering the VRS to this cohort. The predictive validity of the tool and the central risk factors that contribute to recidivism for this cohort remain underexplored.

Lastly, the definition of recidivism used in the study was likely to underestimate the prevalence of further violent behaviour as it was restricted to convictions that resulted in subsequent contact with corrective services. The use of de-identified data meant that it was also not possible to identify convictions that occurred in a different jurisdiction to the one an individual was initially released from.

# Results

## Violence Risk Scale score profile of Australian males

During the study period in which the VRS assessments were conducted, VRS scores were assigned to one of three risk bands: low, moderate or high. Table 1 shows that around half (52.5%) of the 2,689 males recorded a VRS pre-treatment score that fell into the high risk category (a score of greater than 50). The average total VRS score was 50.6 (*SD*=10.8) with a median of 52.0. A chi-square test of independence was performed to examine the relationship between Aboriginality and VRS risk band. The relationship was significant, with a higher proportion of Aboriginal males identified as high risk on the VRS (71.8%) compared with non-Indigenous males (44.7%; $\chi^2(1, n=2{,}689)=173.9$, $p<0.01$). This difference was reflected in their higher average total VRS score of 55.7 (*SD*=9.4) compared with 48.6 (*SD*=10.7).

**Table 1: Distribution of sample by VRS risk category and Aboriginality**

| | Aboriginal | | Non-Indigenous | | Total | |
|---|---|---|---|---|---|---|
| **Risk band** | **Number** | **Percent** | **Number** | **Percent** | **Number** | **Percent** |
| Low | 20 | 2.6 | 221 | 11.5 | 241 | 9.0 |
| Moderate | 198 | 25.6 | 838 | 43.8 | 1,036 | 38.5 |
| High** | 556 | 71.8 | 856 | 44.7 | 1,412 | 52.5 |
| **Total** | **774** | **100.0** | **1,915** | **100.0** | **2,689** | **100.0** |

**statistically significant at $p<0.01$

A recent change to the VRS (Olver et al. 2022) was an increase in the number of risk categories from three to five, along with a move away from the use of stigmatising language such as 'high risk' to describe these risk bands. Table 2 shows the distribution of the sample by Aboriginality based on the revised five-level risk scheme. As shown, non-Indigenous males most frequently fell into risk level III (41.5%) followed by level IV (32.5%), while Aboriginal males most commonly fell into risk level IV (37.9%) and level V (37.1%).

**Table 2: Distribution of sample by five-level risk scheme and Aboriginality**

| | Aboriginal | | Non-Indigenous | | Total | |
|---|---|---|---|---|---|---|
| **Risk level** | **Number** | **Percent** | **Number** | **Percent** | **Number** | **Percent** |
| Level I (0–19.9) | 0 | 0.0 | 10 | 0.5 | 10 | 0.4 |
| Level II (20–34.9) | 16 | 2.1 | 183 | 9.6 | 199 | 7.4 |
| Level III (35–49.9) | 178 | 23.0 | 794 | 41.5 | 972 | 36.1 |
| Level IV (50–59.9) | 293 | 37.9 | 623 | 32.5 | 916 | 34.1 |
| Level V (60–78) | 287 | 37.1 | 305 | 15.9 | 592 | 22.0 |
| **Total** | **774** | **100.0** | **1,915** | **100.0** | **2,689** | **100.0** |

### *Subscale and total VRS scores*

Table 3 presents the scores on the static and dynamic subscales along with the full VRS score pre-treatment. An independent samples *t*-test was conducted to examine whether there were differences in the static and dynamic subtotal scores by Aboriginality. The results indicated that Aboriginal males ($M$=14.1, $SD$=3.0) recorded significantly higher static scores than non-Indigenous males ($M$=11.7, $SD$=3.6; $t(2,875)=18.2$, $p<0.01$). This represented a medium effect size, as measured by Cohen's $d$ ($d$=0.69, 95% CI=[0.61, 0.78]). Aboriginal males also recorded significantly higher dynamic scores ($M$=41.2, $SD$=7.6) compared with non-Indigenous males ($M$=36.6, $SD$=8.5; $t(2,687)=13.6$, $p<0.01$), with a medium effect size ($d$=0.55, 95% CI=[0.47, 0.64]).

**Table 3: Subscale and total VRS scores by Aboriginality**

| | Aboriginal | | | Non-Indigenous | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | *SD* | Median | Mean | *SD* | Median | *d* | 95% CI |
| Static score** | 14.1 | 3.0 | 15.0 | 11.6 | 3.6 | 12.0 | 0.69 | 0.61, 0.78 |
| Dynamic score** | 41.2 | 7.6 | 42.0 | 36.6 | 8.4 | 37.0 | 0.55 | 0.47, 0.64 |
| **Total VRS**** | **55.7** | **9.4** | **57.0** | **48.6** | **10.7** | **49.0** | **0.69** | **0.61, 0.78** |

**statistically significant at $p<0.01$

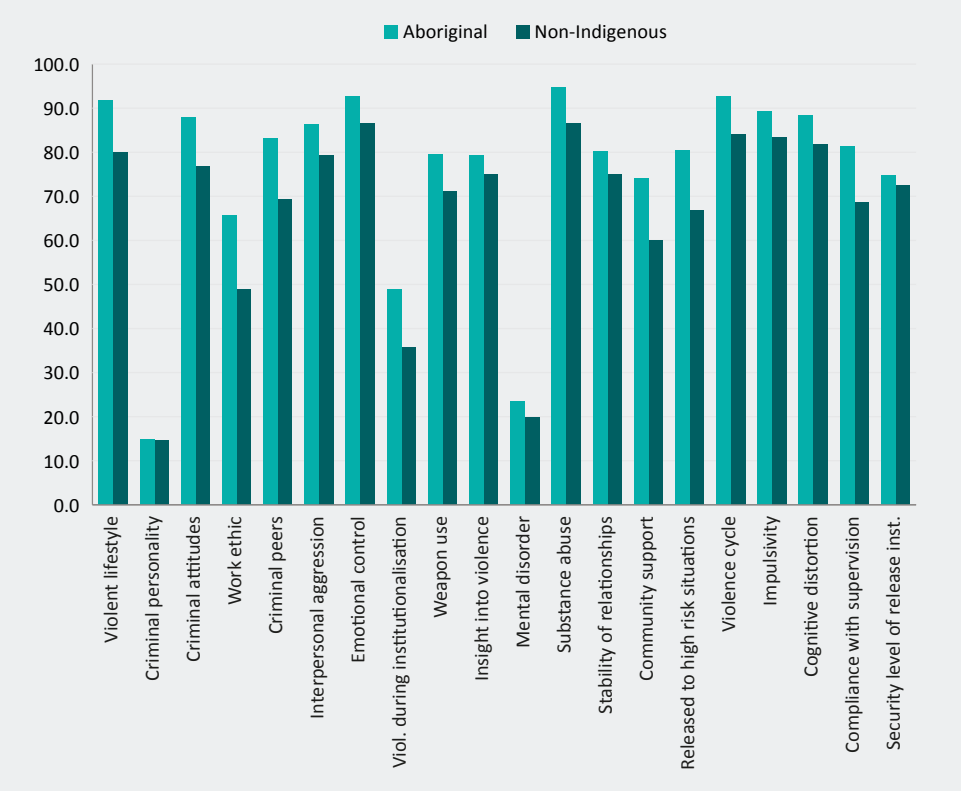
Within the static items, significant differences were noted across all items by Aboriginality excluding current age (at VRS assessment), with Aboriginal males recording consistently higher scores. Most notably, Aboriginal males recorded a higher average risk score for age at first conviction ($M$=2.06), indicating that they were younger at the time of their first conviction than non-Indigenous males ($M$=1.54). An independent samples *t*-test revealed a *t*-statistic of 16.3 ($df$=1,666, $p<0.01$), with a medium effect size ($d$=0.64, 95% CI=[0.56, 0.72]). Aboriginal males also recorded a significantly higher average score for number of youth convictions ($M$=2.27) compared with non-Indigenous males ($M$=1.48; $t(1,746)=15.0$, $p<0.01$). This also had a medium effect size ($d$=0.58, 95% CI=[0.50, 0.66]).

### *Profile of dynamic treatment needs*

Treatment needs for each dynamic item on the VRS were identified for the majority of the sample except for items criminal personality, mental disorder and violence during institutionalization. Emotional control and substance use were identified as needs for most of the sample (88.9% and 88.3% respectively).

Figure 1 presents the proportion of the total sample with a treatment need on each dynamic item by Aboriginality. While the treatment need profile was quite similar overall, there were several areas in which Aboriginal males recorded significantly higher rates of need. This included work ethic, with two-thirds (65.6%) of Aboriginal males recording this item as a treatment target compared with half (48.9%) of non-Indigenous males ($\chi^2(1, n=2,686)=61.8$, $p<0.01$). Community support was also a significantly higher treatment need for Aboriginal males (74.2%) compared with non-Indigenous males (60.0%; $\chi^2(1, n=2,685)=53.2$, $p<0.01$) along with criminal peers, which was a treatment need for 83.2 percent of Aboriginal males compared with 69.4 percent of non-Indigenous males ($\chi^2(1, n=2,680)=47.8$, $p<0.01$). In addition, four-fifths (80.5%) of Aboriginal males recorded a treatment target in relation to release to high-risk situations compared with 66.8 percent of non-Indigenous males ($\chi^2(1, n=2,517)=47.3$, $p<0.01$).

**Figure 1: Proportion of males with a treatment need per dynamic item by Aboriginality (%)**




Note: Dynamic items with missing values were excluded from the percentage calculations

## Recidivism rates for Aboriginal and non-Indigenous males

Table 4 presents the proportion of the sample that reoffended over the two fixed follow-up periods. As shown, one-quarter (25.0%) had been convicted of a violent offence at the two-year mark, with this figure rising to 42.0 percent at five-year follow-up. Aboriginal males recorded consistently higher rates of violent recidivism in each follow-up period than their non-Indigenous counterparts.

**Table 4: Recidivism rates at two-year (*n*=2,125) and five-year (*n*=938) follow-up by Aboriginality**

| | Aboriginal | | Non-Indigenous | | Total | |
|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent |
| Two-year: violent reoffence | 218 | 37.7 | 314 | 20.3 | 532 | 25.0 |
| Five-year: violent reoffence | 146 | 53.9 | 248 | 37.2 | 394 | 42.0 |

Table 5 displays the proportion of Aboriginal and non-Indigenous males who fell into each VRS risk band and the corresponding rate of recidivism by risk category at two-year and five-year follow-up. As expected, those in the low risk band recorded the lowest rate of violent recidivism (15.0% of Aboriginal males and 8.0% of non-Indigenous males at the two-year mark), with reoffending rates increasing with risk level (42.9% of Aboriginal males and 26.8% of non-Indigenous males in the high risk band recorded a violent reoffence within two years).

At five-year follow-up, recidivism rates were relatively comparable for Aboriginal and non-Indigenous males in the moderate risk group (37.0% and 32.5% respectively). For those in the high risk band, there was a similar percentage point difference in the rate of violent recidivism between groups to that observed at two-year follow-up (61.7% for Aboriginal males and 45.5% for non-Indigenous males).

**Table 5: Recidivism rates by VRS risk band and Aboriginality at two-year and five-year follow-up (%)**

| | Aboriginal | | | Non-Indigenous | | |
|---|---|---|---|---|---|---|
| **VRS risk band** | **Low** | **Moderate** | **High** | **Low** | **Moderate** | **High** |
| **Two-year follow-up** | | | | | | |
| Percent in risk band | 3.5 | 26.4 | 70.1 | 11.3 | 47.0 | 41.7 |
| Percent violent reoffence | 15.0 | 26.8 | 42.9 | 8.0 | 17.5 | 26.8 |
| **Five-year follow-up** | | | | | | |
| Percent in risk band | 3.7 | 26.9 | 69.4 | 10.3 | 48.4 | 41.2 |
| Percent violent reoffence | 30.0 | 37.0 | 61.7 | 26.1 | 32.5 | 45.5 |

Table 6 presents two-year and five-year recidivism rates by group based on the updated five-level VRS risk scheme. Consistent with the three risk categories, there was a corresponding increase in the rate of violent recidivism with each risk level. This incremental growth was more pronounced among Aboriginal males in both follow-up periods, whereas non-Indigenous males placed in risk levels IV and V showed similar rates of two-year violent recidivism (25.0% and 29.2%) and the same rate at five-year follow-up (45.0%).

**Table 6: Recidivism rates by five-level risk scheme and Aboriginality at two-year and five-year follow-up (%)**

| | Aboriginal | | | | | Non-Indigenous | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **VRS risk level** | **I** | **II** | **III** | **IV** | **V** | **I** | **II** | **III** | **IV** | **V** |
| **Two-year follow-up** | | | | | | | | | | |
| % in risk band | 0.0 | 2.8 | 23.7 | 39.6 | 34.0 | 0.5 | 9.3 | 44.4 | 31.6 | 14.2 |
| % violent reoffence | – | 6.3 | 25.5 | 37.1 | 49.2 | 0.0 | 6.9 | 17.2 | 25.0 | 29.2 |
| **Five-year follow-up** | | | | | | | | | | |
| % in risk band | 0.0 | 3.0 | 24.0 | 38.4 | 34.7 | 0.4 | 8.1 | 46.2 | 33.3 | 12.0 |
| % violent reoffence | – | 25.0 | 35.4 | 56.7 | 66.0 | 0.0 | 25.9 | 31.8 | 45.0 | 45.0 |

In comparison with the new VRS risk categories and recidivism estimates developed by Olver et al. (2022), the rate of five-year violent recidivism for the total sample in this study was at the upper range or higher than the estimates for most risk levels. This was most notable for those in the level II category, with a violent recidivism rate of 25.8 percent, well above the recommended range of 5 percent to 13.9 percent based on VRS calculator-generated recidivism estimates (Mundt 2020; see Table A1 in the *Appendix*).

### *Predictive validity of VRS scores for violent recidivism*

Relative predictive accuracy of the VRS in discriminating violent recidivists from non-recidivists was assessed using the AUC from ROC analyses. The standardised difference in risk scores between recidivists and non-recidivists was also examined using Cohen's *d*.

The AUCs for the dynamic, static and total VRS scores for violent reoffending are shown in Table 7. All AUCs were in the low to moderate range in their ability to discriminate those who recorded a violent reoffence and those who did not (range of 0.58 to 0.70 across the subscales and VRS total score) and were statistically significant. None of the lower parameters of the 95% confidence intervals were at or below 0.50, indicating that the probability that a randomly selected violent recidivist would have received a higher score on the VRS than a randomly selected non-recidivist was consistently above chance level.

Of note, among Aboriginal males the static score had the highest AUC at two-year (0.66, CI=0.62, 0.71) and five-year follow-up (0.70, CI=0.63, 0.76). This finding did not extend to non-Indigenous males, with that group having similar AUCs in both follow-up periods (0.62 and 0.59 respectively).

**Table 7: Predictive accuracy of the VRS for violent recidivism by Aboriginality**

| | Aboriginal | | | | Non-Indigenous | | | |
|---|---|---|---|---|---|---|---|---|
| | Two-year | | Five-year | | Two-year | | Five-year | |
| | *d* | AUC [95% CI] | *d* | AUC [95% CI] | *d* | AUC [95% CI] | *d* | AUC [95% CI] |
| Static total | 0.53 | 0.66** [0.62, 0.71] | 0.65 | 0.70** [0.63, 0.76] | 0.42 | 0.62** [0.58, 0.65] | 0.34 | 0.59** [0.55, 0.64] |
| Dynamic total | 0.35 | 0.60** [0.56, 0.65] | 0.28 | 0.59* [0.52, 0.65] | 0.35 | 0.60** [0.56, 0.63] | 0.29 | 0.58** [0.54, 0.62] |
| Total VRS score | 0.46 | 0.63** [0.59, 0.68] | 0.43 | 0.63** [0.56, 0.69] | 0.42 | 0.62** [0.58, 0.65] | 0.35 | 0.60** [0.55, 0.64] |

**statistically significant at $p<0.01$, *statistically significant at $p<0.05$

Note: CI=confidence interval

Logistic regression models were used to assess the relationship between total VRS scores and the likelihood of violent recidivism. As shown in Table 8, the analysis yielded an odds ratio of 1.053 for the total sample at two-year follow-up (95% CI=[1.04, 1.06]). This indicates that a 1-point increase in VRS pre-treatment score was associated with a 5 percent increase in the odds of future violence, and this finding remained relatively stable at five-year follow-up (OR=1.045, 95% CI=[1.03, 1.06]). This compares to a slightly higher and increasing odds ratio observed over time by Olver et al. (2022; two-year OR=1.070, 95% CI=[1.042, 1.098]; five-year OR=1.081, 95% CI=[1.060, 1.103]).

**Table 8: Logistic regression models predicting violent recidivism at two-year and five-year follow-up**

| Regression model | *B* | *SE* | Wald | *p* | OR [95% CI] |
|---|---|---|---|---|---|
| **Two-year violent recidivism** | | | | | |
| Total VRS score—Aboriginal | 0.052 | 0.010 | 26.075 | <0.001 | 1.053 [1.032, 1.074] |
| Total VRS score—non-Indigenous | 0.042 | 0.006 | 41.269 | <0.001 | 1.042 [1.029, 1.056] |
| Total VRS score—all | 0.052 | 0.005 | 98.597 | <0.001 | 1.053 [1.043, 1.064] |
| **Five-year violent recidivism** | | | | | |
| Total VRS score—Aboriginal | 0.046 | 0.014 | 11.618 | <0.001 | 1.047 [1.020, 1.075] |
| Total VRS score—non-Indigenous | 0.035 | 0.008 | 18.008 | <0.001 | 1.036 [1.019, 1.053] |
| Total VRS score—all | 0.044 | 0.007 | 41.652 | <0.001 | 1.045 [1.031, 1.059] |

Note: OR=odds ratio; CI=confidence interval

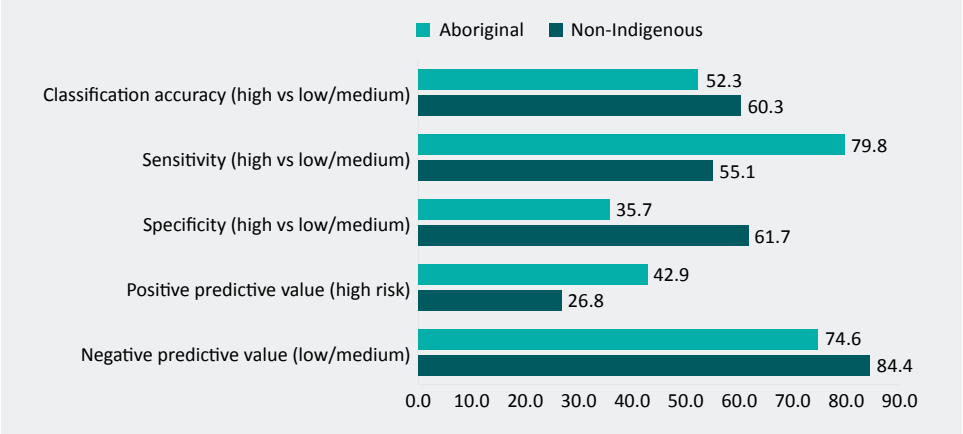
Model statistics at two-year follow-up: Aboriginal—model $\chi^2(df, n)$=28.50 (1, 579), $p$<0.001, Nagelkerke $R^2$=0.07; Non-Aboriginal—model $\chi^2(df, n)$=43.64 (1, 1546), $p$<0.001, Nagelkerke $R^2$=0.04; Total sample—model $\chi^2(df, n)$=108.22 (1, 2125), $p$<0.001, Nagelkerke $R^2$=0.07

Model statistics at five-year follow-up: Aboriginal—model $\chi^2(df, n)$=12.34 (1, 271), $p$<0.001, Nagelkerke $R^2$=0.06; Non-Aboriginal—model $\chi^2(df, n)$=18.81 (1, 667), $p$<0.001, Nagelkerke $R^2$=0.04; Total sample—model $\chi^2(df, n)$=44.55 (1, 938), $p$<0.001, Nagelkerke $R^2$=0.06

Figure 2 presents additional discrimination metrics based on the threshold that meets the classification for high risk on the VRS (a score of >50). Classification accuracy, or the total proportion 'correctly' classified in the high and low/moderate risk groups, was slightly higher for non-Indigenous males (60.3%) than for Aboriginal males (52.3%). In relation to sensitivity, the VRS high risk threshold accurately identified 79.8 percent of Aboriginal males with a violent reoffence. However, specificity was low, with only 35.7 percent of Aboriginal non-recidivists accurately identified below this threshold. In other words, the tool had a high rate of 'false positives', or incorrect classification of Aboriginal males as high risk.

Non-Indigenous males showed the opposite trend. The sensitivity statistic was lower, with 55.1 percent of non-Indigenous violent recidivists correctly classified as recidivists based on the high risk threshold, while specificity was higher, with 61.7 percent of non-recidivists correctly classified below this threshold.

**Figure 2: Additional discrimination metrics of the VRS high risk threshold for two-year violent recidivism by Aboriginality (%)**




Calibration was assessed using the E/O index. This index represents the ratio of the expected number of violent recidivists based on a normative sample to the observed number of violent recidivists in the current study (Hanson 2017). The analyses showed that calibration varied across risk level, with a tendency to underestimate violent recidivism rates in the lower risk categories and overestimate risk in the highest level (see Table A2 in the *Appendix* for E/O index analyses).

## Conclusion

This study examined the profile and discriminative and predictive validity of the VRS in an Australian sample of men convicted of violent offences, including a large number of men of Aboriginal background. Aboriginal men scored higher on the VRS at pre-treatment than their non-Indigenous counterparts, and also recorded significantly higher static and dynamic scores than non-Indigenous men. The total VRS score and subscales were found to have a moderate level of discriminative efficacy for both groups (combined AUC of 0.65). The magnitudes of the AUCs were relatively similar to those reported in the international literature (Olver et al. 2024; Yang, Wong & Coid 2010) and to other risk assessment tools in Australia (Watkins 2011). Furthermore, the findings were broadly consistent with a validation study of the VRS conducted with a sample of high-risk offenders, including a significant number of Canadian First Nations people (Lewis, Olver & Wong 2012).

While the findings of the study indicate the tool has predictive utility with Australian Aboriginal offenders, its relatively modest discriminative accuracy is a clear limitation. Most (72%) of the Aboriginal men in the sample were categorised in the high risk band compared with less than half (45%) of non-Indigenous men. In the context of the tool having only a moderate predictive ability to judge risk of future violence, this has resulted in a disproportionate number of Aboriginal men receiving a 'false positive' rating (ie they were judged to be at high risk but did not go on to commit a violent act). In other words, the tool had lower specificity and higher sensitivity for Aboriginal males than for non-Indigenous males, meaning it was more likely to incorrectly over-classify Aboriginal men as high risk than non-Indigenous men. This is consistent with the findings of a recent Victorian study of the LSI-R (Ashford et al. 2022), which found the scale more likely to classify an Aboriginal non-recidivist as someone who would reoffend than a non-Indigenous person, despite both groups having similar predictive accuracy scores (AUC of 0.60 and 0.63 respectively).

This inaccuracy, along with the much greater volume of Aboriginal men recording very high risk scores, can result in detrimental correctional practices, including more stringent security placements, higher treatment intensity, less likelihood of parole, and even consideration for further supervision orders. Thus, this finding is consistent with the literature that highlights the role that standardised and 'one-size-fits-all' types of risk assessments play in contributing to the over-representation of Aboriginal people in the criminal justice system. As noted by Allan and Dawson (2004), even if risk assessment tools are highly accurate in assessing the risk of recidivism, they may still exclude important predictors of unique relevance to Aboriginal people, a criticism that extends to most psychometric assessment tools (Westerman & Dear 2023). Consequently, their use inhibits our ability to better understand the dynamics of violent offending for this cohort and therefore to develop strategies to reduce future harmful behaviour.

As Skeem and Monahan (2011) noted over a decade ago, the violence risk assessment field may well be reaching a point of diminishing returns in instrument development, as these tools account for very little of the variance in predictive accuracy. Given that an overwhelming need remains to reduce the over-representation of Aboriginal people in the justice system, further research is needed in the short-term to develop culturally relevant semi-structured interview schedules to help practitioners mitigate the risk of differential scoring. Training to build cultural competence is another key component to reducing bias in the risk assessment process. The long-term research agenda, however, should be the co-design of a culturally validated risk assessment scale that can take into account the structural inequalities and disadvantages facing Aboriginal prisoners. Ultimately, correctional agencies in Australia have an inherent responsibility to invest in these areas and encourage ongoing research that will contribute to delivering better life outcomes for Aboriginal and Torres Strait Islander peoples.

## Acknowledgements

This document was developed on Kaurna land. The South Australian Department for Correctional Services (DCS) acknowledges the spiritual connection to Country and recognises the Kaurna people as the custodians of the greater Adelaide region. DCS also recognises the continuing importance of language, land, culture and spiritual beliefs to Aboriginal and Torres Strait Islander people from all regions of Australia, particularly those who will potentially be impacted by this research.

We would like to thank all staff who played a role in the project, particularly the members of the VRS Steering Group and Working Group who made the project possible. Our thanks also extend to Steve Wong and his colleagues for generously sharing their expertise and to the Westerman Jilya Institute for Indigenous Mental Health for its important contribution to this project. We also recognise the contribution of Aboriginal staff from DCS who provided feedback and assistance throughout the project.

## References

ABS—see Australian Bureau of Statistics

Allan A & Dawson D 2004. Assessment of the risk of re-offending by Indigenous male violent and sexual offenders. *Trends & issues in crime and criminal justice* no. 280. Canberra: Australian Institute of Criminology. https://www.aic.gov.au/publications/tandi/tandi280

Andrews DA, Bonta J & Hoge RD 1990. Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior* 17: 19–52. https://doi.org/10.1177/0093854890017001004

Ashford L, Spivak B, Ogloff J & Shepherd S 2022. The cross-cultural fairness of the LS/ RNR: An Australian analysis. *Law and Human Behavior* 46(3): 214–226. https://doi.org/10.1037/lhb0000486

Australian Bureau of Statistics 2022. *Prisoners in Australia*. Canberra: ABS. https://www.abs.gov.au/statistics/people/crime-and-justice/prisoners-australia/2022

Australian Bureau of Statistics 2021. *Estimates of Aboriginal and Torres Strait Islander Australians*. Canberra: ABS. https://www.abs.gov.au/statistics/people/aboriginal-and-torres-strait-islander-peoples/estimates-aboriginal-and-torres-strait-islander-australians/jun-2021

Cohen J 1988. *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum

Gower M, Morgan F & Saunders J 2022. Aboriginality and violence: Gender and cultural differences on the Level of Service/Risk, Need, Responsivity (LS/RNR) and Violence Risk Scale (VRS). *Psychiatry, Psychology and Law* 30(3): 249–270. https://doi.org/10.1080/13218719.2021.2006098

Hanson RK 2017. Assessing the calibration of actuarial risk scales: A primer on the E/O index. *Criminal Justice and Behavior* 44(1): 26–39. https://doi.org/10.1177/0093854816683956

Hsu C, Caputi P & Byrne M 2010. The Level of Service Inventory–Revised (LSI-R): Assessing the risk and need characteristics of Australian Indigenous offenders. *Psychiatry, Psychology and Law* 17: 355–367. https://doi.org/10.1080/13218710903089261

Jones R, Masters M, Griffiths A & Moulday N 2002. Culturally relevant assessment of Indigenous offenders: A literature review. *Australian Psychologist* 37: 187–197. https://doi.org/10.1080/00050060210001706866

Lewis K, Olver ME & Wong SC 2012. The Violence Risk Scale: Predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment* 20(2): 150–164. https://doi.org/10.1177/1073191112441242

Mals P, Howells K, Day A & Hall G 2000. Adapting violence rehabilitation programs for the Australian Aboriginal offender. *Journal of Offender Rehabilitation* 30(1–2): 121–135. https://doi.org/10.1300/J076v30n01_08

Martel J, Brassard R & Jaccoud M 2011. When two worlds collide: Aboriginal risk management in Canadian corrections. *British Journal of Criminology* 51: 235–255. https://doi.org/10.1093/bjc/azr003

Mundt JC 2020. VRS calculator. https://psynergy.ca/vrs

Olver M et al. 2024. Too risky to use, or too risky not to? Lessons learned from over 30 years of research on forensic risk assessment with Indigenous persons. *Psychological Bulletin* 150(5): 487–553. https://doi.org/10.1037/bul0000414

Olver M et al. 2022. Assessing dynamic violence risk: Common language risk levels and recidivism rates for the Violence Risk Scale. *Psychological Assessment* 34(6): 528–545. https://doi.org/10.1037/pas0001116

Productivity Commission 2022. *Closing the Gap annual data compilation report*. Canberra: Commonwealth of Australia. https://www.pc.gov.au/closing-the-gap-data/annual-data-report/2022

Rice ME & Harris GT 2005. Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior* 29: 615–620. https://doi.org/10.1007/s10979-005-6832-7

Shepherd S 2018. Violence risk assessment and Indigenous Australians: A primer. *Alternative Law Journal* 43(1): 45–47. https://doi.org/10.1177/1037969x17748210

Shepherd S, Adams Y, McEntyre E & Walker R 2014. Violence risk assessment in Australian Aboriginal offender populations: A review of the literature. *Psychology, Public Policy, and Law* 20(3): 281–293. https://doi.org/10.1037/law0000017

Shepherd S & Lewis-Fernandez R 2016. Forensic risk assessment and cultural diversity: Contemporary challenges and future directions. *Psychology, Public Policy, and Law* 22(4): 427–438. https://doi.org/10.1037/law0000102

Skeem JL & Monahan J 2011. Current directions in violence risk assessment. *Current Directions in Psychological Science* 20(1): 38–42. https://doi.org/10.1177/0963721410397271

Watkins I 2011. The utility of Level of Service Inventory-Revised (LSI-R) assessments within NSW correctional environments. *Corrective Services NSW Research Bulletin* 29. https://correctiveservices.dcj.nsw.gov.au/resources/research-and-reports/corrections-research-evaluation-and-statistics/research-and-statistical-publications-by-year-of-release.html

Westerman TG & Dear GE 2023. The need for culturally valid psychological assessment tools in Indigenous mental health. *Clinical Psychologist* 27(3): 284–289. https://doi.org/10.1080/13284207.2023.2247532

Woldgabreal Y, Day A, Daffern M, Lloyd C & Graffam J 2022. An empirical test of the factor structure of the Violence Risk Scale and its measurement invariance across time and cultural groups. *Criminal Justice & Behavior* 49(9): 1255–1275. https://doi.org/10.1177/00938548221084984

Woldgabreal Y, Day A & Tamatea A 2020. Do risk assessments play a role in the enduring 'color line'? *Advancing Corrections Journal* 10: 18–28

Wong S & Gordon AE 2006. The validity and reliability of the Violence Risk Scale. A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law* 12(3): 279–309. https://doi.org/10.1037/1076-8971.12.3.279

Wong S & Gordon AE 2003. Violence Risk Scale. https://psynergy.ca/vrs

Yang M, Wong SC & Coid J 2010. The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin* 136(5): 740–767. https://doi.org/10.1037/a0020473

# Appendix

**Table A1: Violent recidivism rates by five-level risk scheme at five-year follow-up (%)**

| | Total | | VRS recommended criterion cut-offs | |
|---|---|---|---|---|
| VRS risk level | Percent in risk band | Percent violent reoffence | Percent in risk band | Recommended range |
| Level I | 0.3 | 0.0 | 10.2 | 0–4.9 |
| Level II | 6.6 | 25.8 | 20.9 | 5–13.9 |
| Level III | 39.8 | 32.4 | 29.9 | 14–31.9 |
| Level IV | 34.8 | 48.8 | 22.9 | 32–49.9 |
| Level V | 18.6 | 56.3 | 16.1 | 50+ |

**Table A2: E/O index and 95% confidence intervals for violent recidivism by five-level risk scheme at five-year follow-up**

| | | Observed Australian sample | | Expected VRS norm | | | |
|---|---|---|---|---|---|---|---|
| VRS risk level | Number in risk category | % violent recidivists | Number violent recidivists | % violent recidivists | Number violent recidivists | E/O index | 95% CI |
| Level I | 3 | 0.0 | 0 | 3.5 | 0.1 | – | – |
| Level II | 62 | 25.8 | 16 | 10.0 | 6.2 | 0.39 | [0.24, 0.63] |
| Level III | 373 | 32.4 | 121 | 26.9 | 100.3 | 0.83 | [0.69, 0.99] |
| Level IV | 326 | 48.8 | 159 | 49.2 | 160.4 | 1.01 | [0.86, 1.18] |
| Level V | 174 | 56.3 | 98 | 63.0 | 109.6 | 1.12 | [0.92, 1.36] |

Note: Values in the 'Expected VRS norm' columns were supplied by Olver et al. (2022) using the VRS calculator algorithm to generate recidivism estimates for the VRS scores in their sample, which were then averaged by risk level. To compute the E/O index, the expected number of violent recidivists was adjusted by applying the expected recidivism rate to the number in each risk category in the current sample

This work was originally published as *Trends & issues in crime and criminal justice* no. 717.
It is available online at **https://doi.org/10.52922/ti77949**.

# 2. Supporting incarcerated mothers: A mixed methods evaluation of the NSW Co-Located Caseworker Program

Althea Gibson, Marc Rémond, Peta MacGillivray, Eileen Baldry and Elizabeth Sullivan

## Introduction

### Mothers in custody in Australia

For several decades, the number of women in custody has been increasing rapidly worldwide (Fair & Walmsley 2022). This global trend has been reflected in Australia, where the female prison population increased by 64 percent between 2009 and 2019 (Australian Institute of Health and Welfare 2020). In the March quarter of 2024, the average daily number of women in custody in Australia was 3,309 (Australian Bureau of Statistics 2024), although this figure does not capture the number of women who enter and exit Australian correctional centres in a relatively short space of time (Baldry 2010).

First Nations women are over-represented in the female prison population in Australia, comprising approximately 45 percent of this cohort (*n*=1,487; Australian Bureau of Statistics 2024). This over-representation can be attributed to a complex interplay of factors, including the ongoing legacy of colonisation, systemic racism, intergenerational and collective trauma, and socio-economic disadvantage (McCausland & Baldry 2023; Newton 2018; Sullivan et al. 2019). Prior involvement in the child protection system is also linked to adult incarceration (Baidawi & Sheehan 2019; Yoorrook Justice Commission 2023), and First Nations children remain over-represented in child protection systems around Australia (Austin 2022; Steering Committee for the Review of Government Service Provision 2024).

A significant proportion of female prisoners in Australia are mothers of minor-aged children, and First Nations women in custody are more likely to be mothers than non-First Nations women (Australian Law Reform Commission 2018). In 2018, the 5th National Prisoner Health Data Collection revealed that 38 percent of prison entrants had children who were reliant on them for their basic needs, with women being more likely than men to report having dependent children (54% vs 36%; Australian Institute of Health and Welfare 2019). In New South Wales, approximately two-thirds of female custodial episodes between 2015 and 2019 (*n*=8,310) involved women with children under 18 years of age, with approximately 60 percent of these children aged 0–10 years, 26 percent aged 11–15 years, and 12 percent aged 16–18 years (Lobo & Howard 2021). Most women involved in these custodial episodes (73%) indicated that their children were not living at home at the time of their entry into custody (Lobo & Howard 2021: 7).

While it is not clear how many mothers in custody have children in contact with child protection services, the limited information available points to a significant intersection between the corrective services and child protection systems. In New South Wales, for example, 34 percent of women entering custody between January 2015 and December 2019 indicated that one or more of their children had care orders at the time of their imprisonment, while 13 percent reported that the Department of Communities and Justice (DCJ; formerly the Department of Family and Community Services) was the care provider for their children (Lobo & Howard 2021).

It is well established that women in prison have gendered needs that have historically been overlooked (Baldry 2010; Bartels, Esteal & Westgate 2020). Relevantly, lack of contact with children is a major source of grief and distress for incarcerated mothers (Baldry 2010; Dworsky et al. 2020; Melander 2020). Contact with children can help reduce the 'isolation, loneliness and helplessness' that many women feel upon separation from their children (United Nations Office on Drugs and Crime 2018: 96), while good mother–child relationships can also help to motivate women to participate in rehabilitative programs (Annie E. Casey Foundation 2011). Children may also benefit greatly from increased mother–child contact, with research demonstrating that the adverse life consequences of having incarcerated parents may be ameliorated by strong mother–child relationships (Arditti & Johnson 2020). Further, and importantly, connection to family members is central to First Nations children's 'sense of identity, belonging and wellbeing' (Davis 2019: 230) and should be supported whenever possible.

While legislation and policy in New South Wales support contact between children in out-of-home care (OOHC) and their families, in practice contact between mothers in custody and their children is often limited (Davis 2019: 328–329). Mothers in custody with children involved in the child protection system may also find it difficult to work towards child protection goals, such as restoration, or to participate in child protection case planning or court proceedings (D'Andrade & Valdez 2012; Halperin & Harris 2004). They may also have unique support needs upon their release from custody. Despite this, services and programs specifically designed for this cohort of women have been largely non-existent.

## The NSW Co-Located Caseworker Program

The Co-Located Caseworker Program (CLCW Program) was introduced in January 2020 to assist women in custody with child protection related issues. A joint initiative of Corrective Services NSW (CSNSW) and Child Protection and Permanency, District and Youth Justice Services (a division of the NSW DCJ), the program 'co-locates' child protection caseworkers in women's correctional centres to assist women with open child protection cases or children in OOHC. Uniquely, the program is available to women on remand as well as sentenced women.

The CLCW Program was designed to contribute to two of the NSW Government's key priorities, as announced by the Premier in 2019: to reduce recidivism in the adult prison population and to increase permanency for children in OOHC. It also responded to recommendations from the NSW *Family is culture* review, which highlighted the need to improve child protection casework practice for Aboriginal children with incarcerated parents (Davis 2019).

After an initial pilot in the Emu Plains Correctional Centre, the program was rolled out to the Silverwater, Dillwynia, Clarence, Mid North Coast and Wellington correctional centres from July 2020. It ceased operating in the Wellington Correctional Centre between June 2021 and November 2022 as the centre was closed due to a mouse plague. It also ceased operating in the Emu Plains Correctional Centre in early 2022, largely due to the closure of the facility during the COVID-19 pandemic.

Co-located caseworker positions were filled by experienced child protection caseworkers employed by the DCJ, including several First Nations caseworkers. Initially, one co-located caseworker was assigned to each correctional centre, although by the time interviews were conducted for this research only one co-located caseworker was working across the Silverwater and Dillwynia correctional centres.

# The evaluation

In 2021, CSNSW commissioned a team of researchers at the University of New South Wales and the University of Newcastle to carry out a mixed methods progressive and summative evaluation of the CLCW Program. The team was asked to address 15 questions about the implementation and effectiveness of the CLCW Program and the experiences of women participating in it, with a specific focus on the experiences of First Nations women (see *Appendix*). Data gathering and analysis were informed by social inclusion theory, as well as rehabilitative, strengths-based, First Nations cultural and women-specific perspectives.

## Research methods

### *Quantitative data sources and linkage*

Records of women who engaged with the CLCW Program between 1 June 2020 and 1 April 2022 were identified from co-located caseworker data, which form part of the DCJ datasets. Records from this dataset were internally linked with the Offender Integrated Management System controlled by CSNSW. Master index numbers were used to link the records from both datasets. After linkage, data were de-identified and supplied to the research team. These data included socio-demographic, criminogenic and health information for each participant in the CLCW Program. In addition, data were obtained relating to the numbers and ages of children to whom the CLCW Program participants were parents or guardians, and these children's care arrangements.

### *Quantitative data analyses*

Descriptive statistics for nominal data are presented as counts and percentages. Continuous measures are presented as mean and standard deviation or median and interquartile range, as appropriate. Women who made up the study cohort were stratified according to two characteristics: First Nations vs non-Indigenous and 'on remand at release' (on remand) vs 'sentenced at release' (sentenced). Differences between these groups were assessed as follows: for continuous variables, the student *t*-test or independent-samples median test were used, as appropriate, while for categorical variables, the chi-square test was used. A $p$-value $<0.05$ was used to infer statistical significance. Data were analysed using SPSS software, version 24 (IBM Corporation, Somers, NY, USA).

### *Qualitative data sources and analyses*

Data were also collected through 48 in-depth interviews with research participants from four groups (see Table 1). Interviews were semi-structured and explored research participants' knowledge and experiences of the CLCW Program, as well as their views about its operation and effectiveness. Women in custody were recruited through a protocol developed by the research team and CSNSW, while other research participants were recruited through a combination of snowball and purposive sampling techniques.

| Table 1: Interview sample | | |
|---|---|---|
| **Interviewees** | **Participant group** | ***n*** |
| Women in custody | Participant group 1 | 25 |
| Women post-release | Participant group 1 | 2 |
| Co-located caseworkers and casework managers | Participant group 2 | 6 |
| CSNSW and child protection staff | Participant group 3 | 9 |
| Non-government/community stakeholders | Participant group 4 | 6 |
| **Total interviews** | | **48** |

Interviews with women in custody took place at the Dillwynia, Silverwater, Clarence and Mid North Coast correctional centres between late August and mid-September 2022. At least one First Nations researcher attended each interview, with approximately half the interviews being conducted by two First Nations researchers. The researchers adopted a motivational interviewing style, which involved taking time to establish rapport with each program participant at the start of each interview, and engaging in active listening throughout the interview to determine which issues program participants were comfortable discussing. While the interviews were semi-structured, program participants were offered the opportunity to 'tell their stories' and subject matter was covered iteratively and organically. A post-interview safety protocol was implemented, whereby one researcher contacted each of the program participants who indicated that they wished to have a wellbeing check after the interview via audiovisual link. During these post-interview conversations, several women indicated that they found the interview experience to have been both empowering and motivating. Of the 25 women in custody interviewed for the evaluation, 11 indicated that they were Aboriginal, while two were from culturally and linguistically diverse backgrounds.

De-identified written transcripts of all the interviews were prepared by the research team and uploaded into the qualitative data analysis software NVivo, version 12. Two researchers conducted a pilot coding process, looking for 'patterns of shared meaning' among the data (Braun & Clarke 2019: 592). A draft coding framework was prepared, and one of the researchers undertook qualitative content analysis of all the transcripts. The approach adopted was one of reflexive thematic analysis and, as such, the draft coding frame was revised and refined as necessary in a flexible data interpretation process.

## The Evaluation Steering Committee

An Evaluation Steering Committee was established to ensure that key stakeholders, including members of Aboriginal-led and community-based organisations, were directly involved in the research process. The Evaluation Steering Committee also helped to ensure that the research team complied with the standards expected in Aboriginal social, emotional and wellbeing research and provided advice about various aspects of the project, including relevant research to draw upon, potential research participants and the framing of the final report.

## Ethical approvals

The primary ethics body for this project was the Aboriginal Health & Medical Research Council's Human Research Ethics Committee (1910/22). Additional ethical approvals were obtained from the University of New South Wales Human Research Ethics Body (HC210987) and the Corrective Services NSW Ethics Committee.

### Limitations

Several variables in the Offender Integrated Management System database are based on self-reported survey responses (ie answers provided by women when completing the Intake Screening Questionnaire upon entering custody) and as such are subject to potential recall issues. Further, as interviews were conducted over a relatively short period, the research team was not able to investigate the longer term views of women in custody about the program or explore the experiences of past program participants. Finally, COVID-19 lockdowns in the research sites affected the delivery of the program, limiting the ability of co-located caseworkers to access correctional centres and engage with women in person, which may have affected women's perceptions of the program.

## Key findings

### Referrals to the program

Analysis of data from CSNSW's Offender Integrated Management System revealed that 400 women were referred to the CLCW Program between 14 June 2020 and 15 March 2022. Referrals were voluntary and were made at the request of the women. Most referrals (54%) were for women in Dillwynia Correctional Centre, while 20 percent were for women in Mid North Coast Correctional Centre, 12 percent for women in Silverwater Women's Correctional Centre, six percent for women in Emu Plains Correctional Centre, six percent for women in Clarence Correctional Centre, and two percent for women in Wellington Correctional Centre.

The caseworkers, casework managers, CSNSW and child protection staff interviewed discussed how the program was promoted when it commenced. Posters were put up in correctional centres to advertise the program and women were informed about it at the regular Inmate Delegate Committee meetings. In some instances, co-located caseworkers also visited groups of women to introduce themselves and the program. Co-located caseworkers described how they generally received from the Services and Programs Officers (SAPOs) the names of women seeking to access the program, although they were sometimes contacted by the Justice Health and Forensic Mental Health Network (Justice Health) about pregnant women in custody. Occasionally, co-located caseworkers were contacted by external child protection caseworkers who had heard about the program and wished to get in contact with mothers in custody. Interviewees, including co-located caseworkers, described how word-of-mouth was a key way in which women found out about the CLCW Program, with one noting that if some women have success with a program, 'a lot of the girls will jump on it'.

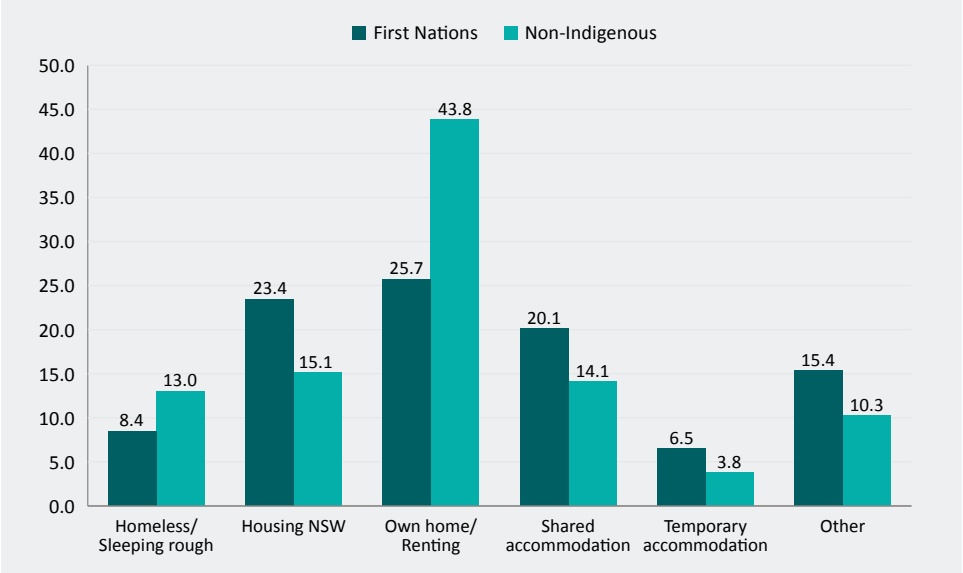
### Participant demographics

Almost all women referred to the program had served a previous custodial episode ($n$=396) and just over one-third (37.8%) were on remand for the duration of their engagement with the program.

Of the 400 women referred to the program, 214 (53.5%) identified as First Nations, while one in five non-Indigenous women were from a culturally and linguistically diverse background. The age of program participants ranged from 18 to 55 years (mean age=32.9 $\pm$ 7.2 years) and First Nations program participants were younger than non-Indigenous program participants (32.2 $\pm$ 7.4 years vs 33.6 $\pm$ 6.9 years, $p$=0.045).

Approximately one-third of women referred to the program (34.0%) reported living in their own home or private rental before entering custody, almost one in five (19.5%) lived in Housing NSW accommodation, and 17.5 percent lived in shared accommodation (see Figure 1). Non-Indigenous women were more likely to live in their own home or private rental before entering custody than First Nations women (43.8% vs 25.7%, $p<0.001$). Forty-two (10.5%) women referred to the program reported that they were homeless or sleeping rough prior to entering custody, and 118 (29.5%) reported that they would need assistance with accommodation after being released from custody.

**Figure 1: Accommodation status of women referred to the CLCW Program, by Indigenous status (%)**




Over one-quarter of women referred to the program (26.5%) reported being the parent/guardian of children under 18 years of age who were living at home, while almost three-quarters (73.5%) were the parent/guardian of children under 18 years of age who were not living at home. Approximately one in seven women (14.5%) reported that their children were in OOHC. Women on remand who were referred to the program were more likely to expect to have custody of their children upon release than women who had been sentenced (32.7% vs 20.2%, $p=0.06$). While a greater proportion of non-Indigenous women (28.4%) expected to have custody of their children after release than First Nations women (22.3%), this difference was not statistically significant ($p=0.166$).

Finally, over two-thirds of women referred to the program (69.5%) reported that they had a mental health condition, while 9.3 percent reported that they lived with disability. Almost one in 12 women also reported having self-harm risk factors.

## Women's experiences of the program

Interviews with women in custody commenced with an exploration of the program's referral process and practices. When asked how they became aware of the program, many women described hearing about it via word of mouth. For some of these women, hearing positive accounts about the co-located caseworkers from other inmates was pivotal to their decision to engage with the program, and this was particularly the case for women who reported feeling distrustful of child protection services generally. For example, one program participant commented:

> One case … she's somebody that didn't want to talk to anybody because of the interaction she's had in the past and it was only after I've given my word and reassured her and I actually came with her to her first meeting and sat with her and given her the confidence … she couldn't thank me enough after it was all over because it was a massive positive outcome for her and it couldn't have gone any better. She didn't expect it to go that well, it was exactly what she needed. And then that's also a person that can vouch to other people. So it's just spreading now …

When asked their views about the program, many women emphatically described how it had benefited them and their children. For the most part, women enjoyed working with the co-located caseworkers, describing them as 'a huge help to everyone', 'amazing', 'very supportive' and 'trustworthy'. Several recounted how the co-located caseworkers advocated for them in a variety of situations, with one commenting that her co-located caseworker helped her to 'have a voice in situations that happen with the kids'. Some women mentioned the positive impact that the program had on the wellbeing of mothers in custody, with one observing that the program helped women 'do our time easier'. Women also expressed a clear preference for face-to-face interactions with co-located caseworkers, indicating that this was important for building rapport and establishing relationships based on trust, while many of the First Nations women interviewed wished to work with First Nations caseworkers.

During interviews, women described how the co-located caseworker had helped them with various child protection related issues. Many were grateful that the co-located caseworker had helped them to locate and re-establish contact with their children. One stated:

> If it wasn't for them I wouldn't have gotten half the way I've gotten. And I'm thankful for that. Because honestly outside I wasn't getting anywhere. And in here I thought that I wouldn't get anywhere. If it wasn't for [co-located caseworkers] I wouldn't have this, I wouldn't be where I am right now getting AVLs [audiovisual link visits].

Other women appreciated receiving assistance to contact external child protection caseworkers, which was described as a difficult and time-consuming process. As one program participant stated:

> … without the DCJ [Department of Communities and Justice] in here, we're constantly running around trying to find an [external] caseworker, trying to find a number … trying to find the time, trying to find the money to make these calls … if there's a middle man then it's easier.

Women also reported receiving updates about the location, placement and wellbeing of their children after engaging with the program, as well as information about court orders, restoration processes and practices, child protection care plans, and their legal rights more generally.

Importantly, some women described how they were able to raise concerns about the safety of their children in OOHC or family placements with the co-located caseworker, who was then able to make enquiries about the wellbeing of their children. Over one-third of the women interviewed for the evaluation indicated that the co-located caseworker had referred them to legal, mental health or other support services.

While all women interviewed for the evaluation agreed that there was a need for the CLCW Program, some expressed concern about delays they experienced accessing the program and dissatisfaction about the length of time between meetings with the co-located caseworker. Some women, including several First Nations women, also expressed frustration at not receiving any updates about their children, noting that some external child protection caseworkers were unwilling to engage with the program. First Nations women also made several suggestions about how to improve the CLCW Program, including delivering the program in group settings and incorporating Elders into the program design.

When discussing their needs and experiences more generally, several women observed that it was difficult to achieve their restoration goals while in custody as they were unable to access programs that addressed their parenting needs and goals. Many women were concerned about their housing on release, with comments including 'Where the hell do I stay?' and 'Where do you stay?'

## Other views about the program

Co-located caseworkers, casework managers, CSNSW and child protection staff who participated in interviews were supportive of the program, unanimously agreeing that it was both necessary and highly successful. These interviewees described the program as 'a great initiative', 'really beneficial', 'crucial', 'absolutely wonderful', 'a valuable program', something that addressed a 'massive gap' and something that 'just works'. Interviewees from CSNSW also observed that the program had improved their workplace efficiency, commenting on the utility of having a single point of contact to deal with child protection matters.

> Previously, you may have spent a good couple of hours, potentially over a number of days, trying to track down what office manages somebody's child … So being able to have in the centre every week a consistent one source of truth is a huge like, I guess, relief. Like resources can be put elsewhere … And it also for the women is reassuring and provides good confidence, so that has been a huge success.

Interviewees from non-government and community organisations also described the CLCW Program as a critical initiative that had the potential to be 'a key driver for change'.

Many interviewees described how the program helped to improve women's motivation and wellbeing. One interviewee commented that mothers in custody:

> … don't see that there's any point to being on the straight and narrow or not going back to the violent ex-partner or not going back to drugs. You know, they're often homeless and they've got no hope of getting their kids back because they can't get a house that they can have the kids in, or the kids are on a track to adoption. And it just feels completely overwhelmingly hopeless … But having that pre-planning while in custody for how everything's going to go when they … get out in relation to their kids is a motivating factor.

Co-located caseworkers confirmed that it was common to see a positive change in the demeanour, outlook and behaviour of women involved in the program, describing that they saw 'boosts' in women's confidence and 'overall kind of positivity'. They also observed that the program motivated women to engage with services to address mental health or drug and alcohol issues and commented that there was a 'calmer environment' in the correctional centres when women had good relationships with their children. In the words of one interviewee:

> … you see the change in them, like it's huge, and even the officers themselves talk about how much of a difference it makes for the inmate within the centre. They might, you know, be violent … causing trouble within the centre and then completely shift and change. … they're happy, they're engaging, [they've] got something to look forward to and they're talking about their child.

Interviewees also described the benefits of the CLCW Program for the children of mothers in custody. Some co-located caseworkers described how they had worked with expectant mothers to organise family placements for babies born in custody, preventing their entry into OOHC. In some cases, when family placements could not be arranged, co-located caseworkers described involving pregnant women in custody in decision-making about their newborn babies. They also observed that through the program they were able to ensure more humane newborn removal practices by, for example, facilitating skin-on-skin contact between mothers and babies and helping post-partum mothers provide their babies with breastmilk.

Further, several interviewees described the profound impact of increased contact between mothers and their children on children's emotional and mental wellbeing. One interviewee observed that contact between a child and mother:

> … could be all the difference in terms of a placement breaking down, to a child engaging well at school, to … Aboriginal kids knowing who their family is, knowing which country they come from. All of these things are absolutely imperative to ensure the best outcomes for kids.

Interviews revealed the co-located caseworkers often carried active caseloads of between 20 and 60 women at any one time. One co-located caseworker commented that it was generally possible for her to see between five and 10 women a day.

## Discussion

Women in custody with children involved in the child protection system represent a substantial yet largely overlooked subset of the women's prison population. Traditionally, these women have been required to simultaneously engage with the corrective services and child protection systems—two large bureaucracies with vastly different aims, policies, practices and organisational cultures. This can pose a significant challenge for women attempting to organise contact with their children, access prison programs that address child protection concerns, or plan for re-integration into their children's lives upon their release. Introduced in 2020, the CLCW Program was designed to ameliorate some of these challenges by increasing cross-system collaboration between CSNSW and the child protection division of DCJ in New South Wales.

The evaluation of the CLCW Program revealed that it was well received by women in custody, as evidenced by high self-referral rates and the positive views of program participants. The program was also seen as an extremely beneficial initiative by CSNSW and child protection staff, and non-government organisations. There were no difficulties in the implementation of the program, which had significant stakeholder 'buy-in' from the outset. Co-located caseworkers did, however, identify some aspects of the correctional environment that could be addressed to improve their efficiency, such as the lack of access to mobile telephones and correctional services data systems. They also observed that addressing impediments to women in custody maintaining contact with their children, including the high cost and short duration of telephone calls, would help women stay connected with their children during periods of incarceration.

The over-representation of First Nations women in custody in Australia is widely documented (Australian Bureau of Statistics 2024), and over half the 400 women referred to the CLCW Program during the study period identified as First Nations. Further, 11 of the 25 women interviewed for the evaluation were First Nations, with many expressing a clear preference for working with First Nations child protection caseworkers. As such, the inclusion of First Nations co-located caseworkers helped to ensure program acceptance and accommodated some of the cultural needs of program participants. There is scope, however, to redesign the program to include more culturally connected elements. For example, as suggested in interviews, the program could involve Elders to encourage, support and motivate First Nations mothers facing adversity.

The evaluation revealed that the co-located caseworkers performed a wide range of activities that far exceeded initial expectations of their roles and responsibilities. This demonstrates the diversity and complexity of the child protection related needs of women in custody, a cohort already widely recognised as having complex support needs arising from intersecting experiences of poverty, domestic and family violence, sexual abuse, housing instability, substance abuse, childhood trauma, and mental health disorders and/or cognitive disability (Australian Institute of Health and Welfare 2020; Baldry 2010; Breuer et al. 2021; Day et al. 2018; McCausland & Baldry 2023, 2017).

Under the program, co-located caseworkers acted as a point of contact and source of information for both women in custody and external child protection caseworkers. In this unique role, they were able to complete tasks that incarcerated women and CSNSW staff found difficult, such as identifying and contacting external child protection caseworkers and ascertaining the location of children in OOHC placements. They were also able to assist external child protection caseworkers to locate and contact women in custody, including by supporting these caseworkers with practical issues such as booking audiovisual link visits or telephone calls with women in custody.

Importantly, co-located caseworkers worked with Justice Health to engage in prenatal casework with pregnant women in custody. While there is limited information about rates of pregnancy among Australian women in custody, 5.7 percent of Indigenous and 4.0 percent of non-Indigenous women admitted into custody between 2001 and 2011 in Western Australia were pregnant at the time of their incarceration (Dowell et al. 2018). Expectant mothers in custody have unique health and wellbeing needs and are often 'at the extreme end of the continuum of social disadvantage' (Walker et al. 2014: 8). Despite this, child protection casework with women in custody is limited, with many not being allocated a child protection caseworker and some only informed about whether they are able to access mother and child programs with their infants immediately prior to or at the time of birth (Bartels, Esteal & Westgate 2020; Davis 2019). Interviewees noted that 'four or five' babies born to women in custody had not been placed in OOHC as a result of the casework undertaken by the co-located caseworkers. In these cases, the co-located caseworkers had worked with mothers to find their newborn babies suitable placements until the mothers' release from custody. In cases where babies had to go into care due to a lack of placement options, the work of the co-located caseworkers helped to reduce the shock and trauma associated with newborn removals (Davis 2019).

The evaluation concluded that the CLCW Program has potential important long-term benefits for women in custody. First, by increasing mother–child contact, the program has the potential to greatly improve the wellbeing, motivation and behaviour of women in custody. Data collected during interviews for the evaluation were consistent with these research findings, with several women in custody describing the severe anxiety they experienced after being separated from their children and other interviewees commenting on the noticeable positive changes in the women involved in the program.

The evaluation also demonstrated that the CLCW Program could have important benefits for the children of women in custody. In addition to the significant psychosocial benefits to children of increased contact with their mothers, the evaluation revealed that the CLCW Program provided an important mechanism for women in custody to raise concerns about the safety of the children in family or OOHC placements.

Although the evaluation was not able to investigate the long-term impact of the CLCW program on whether children in OOHC were eventually returned to their mothers, the program has the potential for positive restoration impact. Women in custody described how they had obtained information about restoration practices and processes from their co-located caseworker, and other interviewees described how several mothers involved in the program were on track to achieve restoration of their children post-release. In addition, co-located caseworkers regularly referred women in custody to external legal and support services, which could help women address barriers to restoration.

However, some women in custody were concerned about delays in having their referrals to the CLCW Program accepted and delays between caseworker meetings, highlighting the need to ensure there are enough co-located caseworkers to engage in thorough and careful casework with all women referred to the program. Many women in custody also expressed frustration at their inability to access the programs that child protection caseworkers had advised them to complete. There was also no indication that child protection caseworkers considered the adequacy of the CSNSW programs undertaken by women in custody during child protection case planning. The evaluation concluded that there was scope for improved coordination between CSNSW and the child protection division of DCJ with respect to case planning to ensure that women in custody could access child protection related programs. These programs should be trauma- and culturally-informed, and developed and offered within a gender-responsive framework (Bartels, Esteal & Westgate 2020; Bevis et al. 2020).

The evaluation also determined that greater involvement of the NSW Government's housing service (DCJ Housing) in the CLCW Program could result in better outcomes for women and their children. It has long been recognised that women in custody struggle to access adequate and affordable housing upon their release (see, for example, Baldry & McCausland 2009; Baldry et al. 2003). The data analysed for the evaluation was consistent with this research, revealing that 10.5 percent of women referred to the program were homeless or sleeping rough prior to entering custody, and almost one-third said that they would need assistance with accommodation upon their release. Given that a lack of adequate and affordable housing options is a barrier to achieving child protection goals, the CLCW Program would benefit from the inclusion of representatives from DCJ Housing in the future.

## Conclusion

The independent progressive and summative evaluation of the CLCW Program in New South Wales concluded that it was a well designed, effectively implemented and much needed initiative for women in custody who have children involved with the child protection system. All participants involved in the program were enthusiastic about the outcomes it had achieved for women and children in a short period and the possibility of its expansion and refinement in the future. The program demonstrates the potential to have far-reaching positive benefits for women in custody and children involved in the child protection system. It also has the potential to improve interagency coordination and collaboration between CSNSW and the child protection division of DCJ in New South Wales. The program may be improved in several ways if expanded, including by greater involvement of other services such as DCJ Housing and increased emphasis on the coordination of correctional and child protection case planning.

## Funding

This research was funded by Corrective Services NSW, a division of the Department of Communities and Justice.

## Acknowledgements

The authors would like to thank everyone who volunteered to participate in or assist with this research, including the women in custody who volunteered to be interviewed and the members of the Evaluation Steering Committee. We also thank Kelly Austin from CSNSW and Mark Howard from Corrections Research Evaluation and Statistics for their ongoing assistance with the evaluation.

## References

*URLs correct as of October 2024*

Annie E. Casey Foundation 2011. *When a parent is incarcerated: A primer for social workers*. https://www.aecf.org/resources/when-a-parent-is-incarcerated

Arditti J & Johnson EI 2020. A family resilience agenda for understanding and responding to parental incarceration. *American Psychologist* 77(1): 56–70. https://doi.org/10.1037/amp0000687

Austin S 2022. 'Our kids belong with family': A look into institutional child removal. IndigenousX. https://indigenousx.com.au/our-kids-belong-with-family-a-look-into-institutional-child-removal/

Australian Bureau of Statistics 2024. Corrective Services, Australia: National and state information about adult prisoners and community-based corrections, including legal status, custody type, Indigenous status, sex. Reference period: March Quarter 2024. https://www.abs.gov.au/statistics/people/crime-and-justice/corrective-services-australia/latest-release

Australian Institute of Health and Welfare 2020. *The health and welfare of women in Australia's prisons*. AIHW cat. no. PHE 281. Canberra: AIHW. https://www.aihw.gov.au/reports/prisoners/health-and-welfare-of-women-in-prison/summary

Australian Institute of Health and Welfare 2019. *The health of Australia's prisoners 2018*. AIHW cat. no. PHE 246. Canberra: AIHW. https://doi.org/10.25816/5ec5c381ed17a

Australian Law Reform Commission 2018. Pathways to Justice: An inquiry into the incarceration rate of Aboriginal and Torres Strait Islander peoples. ALRC Report 133. Sydney: Commonwealth of Australia. https://www.alrc.gov.au/publication/pathways-to-justice-inquiry-into-the-incarceration-rate-of-aboriginal-and-torres-strait-islander-peoples-alrc-report-133/

Baldry E 2010. Women in transition: From prison to … . *Current Issues in Criminal Justice* 22(2): 253–267. https://doi.org/10.1080/10345329.2010.12035885

Baldry E & McCausland R 2009. Mother seeking safe home: Aboriginal women post-release. *Current Issues in Criminal Justice* 21: 288–300. https://doi.org/10.1080/10345329.2009.12035846

Baldry E, McDonnel D, Maplestone P & Peeters M 2003. *Ex-prisoners and accommodation: What bearing do different forms of housing have on social reintegration: Final report*. Report no. 46. Melbourne: Australian Housing and Urban Research Institute. https://www.ahuri.edu.au/research/final-reports/46

Baidawi S & Sheehan R 2019. 'Crossover kids': Offending by child protection-involved youth. *Trends & issues in crime and criminal justice* no. 582. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti04138

Bartels L, Esteal P & Westgate R 2020. Understanding women's imprisonment in Australia. *Women & Criminal Justice* 30(3): 204–219. https://doi.org/10.1080/08974454.2019.1657550

Bevis M, Atkinson J, McCarthy L & Sweet M 2020. *Kungas' trauma experiences and effects on behaviour in Central Australia*. Research Report 03/2020. Sydney: Australia's National Research Organisation for Women's Safety. https://www.anrows.org.au/project/kungas-trauma-experiences-and-effects-on-behaviour-in-central-australia/

Braun V & Clarke V 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11(4): 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Breuer E, Remond M, Lighton S, Passalaqua J, Galouzis J, Stewart KA & Sullivan E 2021. The needs and experiences of mothers while in prison and post-release: A rapid review and thematic synthesis. *Health & Justice* 9(31). https://doi.org/10.1186/s40352-021-00153-7

D'Andrade AC & Valdez M 2012. Reunifying from behind bars: A quantitative study of the relationship between parental incarceration, service use and foster case reunification. *Social Work in Public Health* 27(6): 616–636. https://doi.org/10.1080/19371918.2012.713294

Davis M 2019. *Family is culture: Independent review of Aboriginal children and young people in OOHC*. https://dcj.nsw.gov.au/children-and-families/family-is-culture/independent-review/media/independent-review-of-aboriginal-children-young-people-in-oohc.html

Day A, Casey S, Gerace A, Oster C & O'Kane D 2018. *The forgotten victims: Prisoner experience of victimisation and engagement with the criminal justice system*. Research Report 01/2018. Sydney: Australia's National Research Organisation for Women's Safety. https://www.anrows.org.au/publication/the-forgotten-victims-prisoner-experience-of-victimisation-and-engagement-with-the-criminal-justice-system

Dowell CM, Mejia GC, Preen DB & Segal L 2018. Maternal incarceration, child protection and infant mortality: A descriptive study of infant children of women prisoners in Western Australia. *Health and Justice* 6(2). https://doi.org/10.1186/s40352-018-0060-y

Dworsky A, Fedock G, Schlecht C, Malcome M, Murray C & Hazel C 2020. Addressing the needs of incarcerated mothers and their children in Illinois. Chicago: Chapin Hall (University of Chicago). https://www.chapinhall.org/wp-content/uploads/PDF/Incarcerated-Mothers-FINAL-2020.pdf

Fair H & Walmsley R 2022. World Female Imprisonment List (5th ed). London: Institute for Crime & Justice Police Research, Birkbeck University. https://www.icpr.org.uk/news-events/2022/world-female-prison-population-60-2000

Halperin R & Harris JL 2004. Parental rights of incarcerated mothers with children in foster care: A policy vacuum. *Feminist Studies* 30(2): 339–352. https://doi.org/10.2307/20458967

Lobo J & Howard M 2021. *Women in prison: An examination of the support needs of women in custody with children*. Research Brief No 4. Sydney: Corrective Services NSW

McCausland R & Baldry E 2023. Who does Australia lock up? The social determinants of justice. *International Journal for Crime, Justice and Social Democracy* 12: 37–53. https://doi.org/10.5204/ijcjsd.2504

McCausland R & Baldry E 2017. Understanding women offenders in prison. In J Ireland, C Ireland, M Fisher & N Gredecky (eds), *The Routledge international handbook of forensic psychology in secure settings*. London: Routledge: 25–39. https://doi.org/10.4324/9781315673073

Melander L 2020. Keeping connected while living apart: An exploration of a prison contact maintenance program. *Journal of Offender Rehabilitation* 59(5): 267–284. https://doi.org/10.1080/10509674.2020.1745980

Newton BJ 2019. Understanding child neglect in Aboriginal families and communities in the context of trauma. *Child & Family Social Work* 24: 218–226. https://doi.org/10.1111/cfs.12606

Steering Committee for the Review of Government Service Provision 2024. *Report on government services 2024*. https://www.pc.gov.au/ongoing/report-on-government-services/2024

Sullivan EA et al. 2019. Aboriginal mothers in prison in Australia: A study of social, emotional and physical wellbeing. *Australian and New Zealand Journal of Public Health* 43(3): 241–247. https://doi.org/10.1111/1753-6405.12892

United Nations Office on Drugs and Crime 2018. *Introductory handbook on the prevention of recidivism and the social reintegration of offenders*. Criminal Justice Handbook Series. Vienna: United Nations. https://www.unodc.org/documents/justice-and-prison-reform/18-02303_ebook.pdf

Walker JR, Hilder L, Levy MH & Sullivan EA 2014. Pregnancy, prison and perinatal outcomes in New South Wales, Australia: A retrospective cohort study using linked health data. *BMC Pregnancy & Childbirth* 14: 214. https://doi.org/10.1186/1471-2393-14-214

Yoorrook Justice Commission 2023. *Report into Victoria's child protection and criminal justice systems*. Melbourne: Yoorrook Justice Commission. https://yoorrookjusticecommission.org.au/topic/yoorrook-for-justice/

# Appendix

| Table A1: Evaluation questions |
|---|
| **Program implementation and effectiveness** |
| Has the program been implemented as intended? |
| What is the journey of women who participate in this initiative? |
| What are staff (CSNSW and caseworker) perceptions of the service? |
| How does the case management intersect with other needs of women who are parents, such as housing, health etc? Does the role assist women navigate the system generally? |
| Is the program effective in connecting women to community-based services once they leave custody? |
| **Women's experiences** |
| How do the women perceive the co-located case managers? |
| What do women want from this service? |
| How would women consider success of this program? |
| What have you received as part of the case management service? |
| What knowledge have you gained as part of this service? |
| What have been the benefits to you as a mother from this service? |
| What has been the impact on the interaction with the justice system/CSNSW? |
| **Aboriginal women's experiences** |
| Have you been able to access cultural services through the case management service? |
| Were these services effective? |
| Were the services culturally safe? |

This work was originally published as *Trends & issues in crime and criminal justice* no. 709. It is available online at **https://doi.org/10.52922/ti77741**.

# 3. The extent of illicit drug and alcohol involvement in crime: An updated estimate

Alexandra Voce, Tom Sullivan, Anthony Morgan and Merran McAlister

There is a well-documented connection between illicit drug and alcohol misuse and criminal offending (Bennett, Holloway & Farrington 2008; Duke et al. 2018). While most people in the general population who use illicit drugs and alcohol do not engage in criminal activity, except where they commit offences related to drug possession or supply, substance use is one among many complex and interrelated factors that can contribute to criminal offending (Zhong, Yu & Fazel 2020). Intoxication can impair judgement and lower inhibition, which increases the risk of intoxicated individuals engaging in criminal behaviours, such as physical violence or public disorder (Darke 2010; McKetin et al. 2014). Some people who use substances may also fund their substance use by committing acquisitive offences, such as theft (Sutherland et al. 2023), particularly if they are dependent on the substance (Bennett, Holloway & Farrington 2008). The connection between substance use and crime varies depending on the substance, the frequency of use and amount consumed, the category of offence, and personal characteristics such as age and sex (Bennett, Holloway & Farrington 2008; Foran & O'Leary 2008; Nagar & Rabinovitz 2024; Zhong, Yu & Fazel 2020).

Understanding the complex interplay between substance use and criminal activity is important in informing law enforcement and health strategies aimed at reducing drug-related crime, allowing policymakers to allocate resources effectively, and providing estimates against which national and jurisdictional policies can be compared and evaluated (Payne & Gaffney 2012). Moreover, these data are important to estimate the costs of drug- and alcohol-related crime to the Australian community (National Drug Research Institute 2020). A lack of contemporary Australian evidence on the extent to which crime is attributable to drug or alcohol use was highlighted in a recent inquiry held by the Parliamentary Joint Committee on Law Enforcement (2023). To gain a better understanding of the nature and extent of drug-related offending, the committee recommended that further research be undertaken.

The extent to which substance use contributes to offending has been measured using data collected by the Drug Use Monitoring in Australia (DUMA) program. The national DUMA program operated from 1999 to 2021, and involved interviewing police detainees about their drug use, including whether their use of different substances contributed to their offending. Based on DUMA data collected in 2009, Payne and Gaffney (2012) found that nearly half of all police detainees (45%) attributed their current offending to substance use, and that offending was more frequently attributed to alcohol than to all other drugs combined. These findings were consistent with self-report data from 29,138 Canadian offenders surveyed between 2006 and 2016, 42 percent of whom said their offending would not have occurred if they were not seeking or under the influence of alcohol or drugs (Young et al. 2021). Similar results have been found in the United States (National Drug Intelligence Center 2011). Unsurprisingly, research has shown the extent of alcohol and drug involvement varies by crime type (Håkansson & Jesionowska 2018). Violent crime, for instance, is more strongly associated with alcohol and stimulants (such as methamphetamine) than with sedatives or opiates (such as heroin or benzodiazepines).

The proportion of offences attributable to drugs or alcohol (hereafter called the 'attributable fraction') calculated using DUMA data has also been applied to crime statistics released by the Australian Bureau of Statistics (ABS). This produces an estimate of the total number of criminal charges in Australia that are attributable to drugs or alcohol. This approach was demonstrated by Morgan and colleagues (2018), who used data on the involvement of alcohol in violent offending as reported by DUMA detainees (35%) to estimate that 24,798 offenders proceeded against for acts intended to cause injury in 2013–14 attributed their offending to alcohol consumption. This estimate was limited by inconsistencies between definitions used in DUMA and by the ABS, and it is unknown how many more criminal offences would be attributable to alcohol and other drug use if this method were applied to all categories of crime, rather than only acts intended to cause injury.

Further, since the earlier study by Payne and Gaffney (2012), the prevalence and patterns of substance use among Australians have changed. Recent data (Australian Institute of Health and Welfare (AIHW) 2025a; Voce & Sullivan 2022a) suggest alcohol consumption has decreased over the past decade, while consumption of specific illicit drugs, such as methamphetamine and cocaine, has increased. It is not clear whether the relationship between alcohol and other drug use and crime has also changed over this period.

To better understand the contemporary relationship between drugs and crime in Australia, this study used data from the DUMA program to measure the extent to which police detainees in Australia attributed their offending to illicit drug and alcohol use, updating the results from the Payne and Gaffney (2012) study. These attributable fractions were then used to estimate the total number of recorded offences attributable to illicit drug and alcohol use.

## Methodology

This study uses data from the DUMA program collected in 2019. These were the most recent data that were not impacted by the COVID-19 pandemic, which was associated with significant short-term changes in alcohol and drug use (AIHW 2025a; Doherty, Sullivan & Voce 2021; Voce, Sullivan & Doherty 2021) and offending (Payne, Morgan & Piquero 2022) among Australians. Wastewater analyses (Australian Criminal Intelligence Commission (ACIC) 2024) and self-report survey data (AIHW 2025b) show that most Australian illicit drug markets had returned to pre-pandemic levels by mid-2022. Likewise, self-reported alcohol consumption in Australia has remained stable or increased since the pandemic (AIHW 2025b). Nonetheless, the use of 2019 data is a limitation to be considered.

The DUMA program collected drug use, criminal offending, and sociodemographic information from police detainees at watch houses or police stations. Participation in the program was voluntary, anonymous and confidential. In 2019, DUMA collected survey data from 2,330 adult detainees (aged 18 years or over) at five DUMA sites in Adelaide, Brisbane, Perth, and Bankstown and Surry Hills in Sydney. For the current study, detainees were excluded if they answered 'don't know' or declined to answer any of the relevant questions regarding past-month drug and alcohol use ($n$=31) or questions regarding the substance attributions ($n$=50). Of the total remaining sample of 2,249 detainees, 81 percent were male ($n$=1,815), three-quarters (74%, $n$=1,660) self-identified as non-Indigenous, and the median age was 33 (interquartile range=26–41).

The survey asked detainees if they had consumed alcohol or used cannabis, methamphetamine (or speed/ice), heroin or ecstasy (or MDMA) in the past month. For each substance a detainee had used, the survey asked: 'Thinking again about the main reason why you have been detained today … how much do you think (alcohol/cannabis/heroin/methamphetamine/ecstasy) contributed to what happened?' Detainees could respond 'not at all', 'a little' or 'a lot'. Consistent with previous research (National Drug Research Institute 2020; Payne & Gaffney 2012), we defined detainees who responded either 'a little' or 'a lot' as attributing their current offending (ie offences for which they were detained) to substance use. For each illicit drug a detainee had used, the survey then asked about the reasons why their drug use had contributed to their offending: 'How would you describe the role of (cannabis/heroin/methamphetamine/ecstasy) in what happened?' Detainees could select one or more of the reasons listed, which included that they needed money to buy the drug, they were 'high' on the drug at the time, they were 'hanging out' for the drug, and an 'other' reason related to their use of that drug.

The relationship between substance use and crime is expressed as attributable fractions—the proportion of all detainees, or detainees charged with a specific offence type, who attributed their offending to alcohol or illicit drugs. We used police charge records to obtain offence information for detainees who completed the questionnaire. A maximum of 10 charges were recorded, all of which related to the detainee's current detention. Charges were assigned to each detainee (including their most serious offence type) according to the third edition of the Australian and New Zealand Standard Offence Classification (ANZSOC; ABS 2011). Because detainees were asked about the contribution of alcohol or drugs to the *main* reason they were detained, we used their principal offence—the most serious charge according to the ABS National Offence Index—when determining whether substance use had contributed to their offending (ABS 2023).

To calculate the approximate number of criminal offences attributable to drug and alcohol use during the same year, we used ABS data on the number of alleged offenders aged 18 years or over who were proceeded against by state and territory police between 1 July 2018 and 30 June 2019 (ABS 2024). The data report the principal offence (most serious offence) allegedly committed by an individual offender based on the ABS National Offence Index (ABS 2023). Offenders who were proceeded against more than once in the reference period (repeat offenders) were assigned a principal offence for each separate date of police action. Using principal offence as the unit of measurement means the data do not reflect the total number of individual offences that come to the attention of police. However, it is more accurate to rely on data relating to offenders than incidents or victims, given the attributable fractions are drawn from interviews with detainees charged with an offence. We nevertheless acknowledge that there are limitations regarding the direct comparability of data on detainees charged with an offence and data on offenders proceeded against by police.

## Limitations

Several methodological limitations to attributable fraction data from the DUMA program must be noted from the outset. The program surveyed adult offenders from select police stations and watch houses in inner-city metropolitan sites, meaning that the attributable fractions may not generalise to the broader offender population in Australia, particularly those living in regional or remote areas (Voce & Sullivan 2022a). Data were not collected from every state and territory (with the Northern Territory, Tasmania, Victoria and the Australian Capital Territory omitted). The ACIC's National Wastewater Drug Monitoring Program (ACIC 2024) has shown that consumption of major drugs was generally comparable across the states included and excluded from DUMA, particularly in terms of which substances are used more and less frequently. However, it also demonstrates that regional variability exists.

The DUMA program also excluded detainees who did not consent to participate or who could not be surveyed due to intoxication, violent behaviour or other reasons—and they may have been more or less likely to attribute their offending to substance use (Voce & Sullivan 2022a). The data also rely on detainees' self-reported attributions and it is not known whether any systematic biases may influence the degree to which offenders attribute their offending to substance use. For example, some offenders may believe that exaggerating the influence of drugs or alcohol on their criminal activity may reduce their culpability (National Drug Research Institute 2020; Payne & Gaffney 2012). Attribution questions were only asked of detainees who had recently used drugs. Thus, the attributable fractions do not capture those detainees who do not regularly use drugs but whose offending may have nonetheless been influenced by the illicit drug market—for example, drug dealers who engage in violence to recover debts for their buyers (National Drug Research Institute 2020). Moreover, the use of self-report drug attributions among apprehended offenders, along with recorded crime data, means that the current estimates do not account for offences that did not result in an offender being proceeded against or which were not reported to or recorded by police.

Several methodological differences should be noted between the current and previous study of DUMA detainees (Payne & Gaffney 2012). Detainees in 2009 were recruited from nine (rather than five) data collection sites around Australia, including sites in Victoria and the Northern Territory which were not represented in the 2019 survey. In 2009, detainees were asked attribution questions about a much broader range of drugs than in 2019, including illicit opiates, benzodiazepines, inhalants and hallucinogens. These differences should be taken into consideration when comparing the findings between the two studies. For example, we are unable to compare results for all drugs or all substances, and limit any comparisons to those substances that were measured in both study periods. In addition, newer synthetic opioids were not measured in either study but remain of concern (ACIC 2024; Brown & Morgan 2019). Cocaine was also excluded. It is one of the more commonly used illicit drugs in Australia but did not feature prominently among police detainees. Further research is needed to account for the possible role that these drugs might play in offending.

# Results

## Attributions by substance type

Two-thirds of the sample (*n*=1,511, 67.2%) reported using at least one illicit drug in the past month and 56.5 percent (*n*=1,270) reported drinking alcohol (see Table 1). In total, 84.7 percent (*n*=1,905) of detainees reported some substance use (drugs or alcohol) in the past month. Of detainees who had used in the past month, 53.3 percent (*n*=1,016) reported that their substance use had contributed to their current offending. This is equivalent to 45.2 percent of all detainees. Twenty-nine percent (*n*=370) of past-month alcohol users reported that alcohol had contributed to their offending (16.5 percent of all detainees), and 49.4 percent (*n*=747) of past-month illicit drug users reported that drugs had contributed to their offending (33.2 percent of all detainees). Over half of all detainees who had used methamphetamine in the past month (57.8%, *n*=620) attributed their offending to methamphetamine, the highest attribution rate among all substances analysed. This was equivalent to 27.6 percent of all detainees.

**Table 1: Past-month drug use and attribution by substance type**

| | Used substance in the past month[a] | | Detainees who attributed offending to substance use | | |
|---|---|---|---|---|---|
| | *n* | % | *n* | % of past-month users[b] | % of total sample |
| Alcohol | 1,270 | 56.5 | 370 | 29.1 | 16.5 |
| Cannabis | 1,041 | 46.3 | 155 | 14.9 | 6.9 |
| Heroin | 164 | 7.3 | 64 | 39.0 | 2.9 |
| Methamphetamine | 1,073 | 47.7 | 620 | 57.8 | 27.6 |
| Ecstasy | 93 | 4.1 | 11 | 11.8 | 0.5 |
| Any illicit drug | 1,511 | 67.2 | 747 | 49.4 | 33.2 |
| Any substance | 1,905 | 84.7 | 1,016 | 53.3 | 45.2 |

a: Calculated as a percentage of all detainees whose responses to the relevant past-month use and attribution questions were not missing or 'don't know'

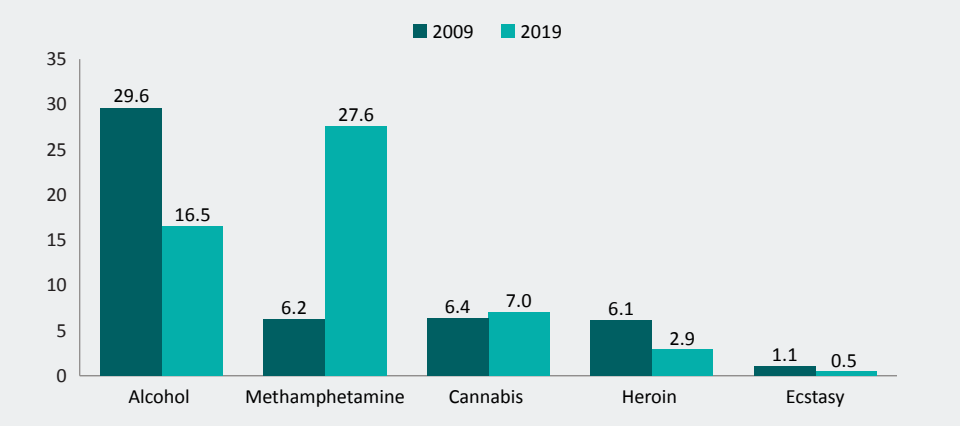
b: Calculated as a percentage of detainees who used a substance in the past month and whose response to the attribution question was not missing or 'don't know'

Note: Total sample size varies for each substance due to missing data. Includes detainees who did and did not have charge data

Source: AIC DUMA collection 2019 [computer file]

We compared the results from the current study with those from the Payne and Gaffney (2012) study. There were several key differences. Most notably, the proportion of all detainees who attributed their offending to alcohol nearly halved, falling from 29.6 percent to 16.5 percent. The proportion of detainees who attributed their offending to heroin also decreased, from 6.1 percent in 2009 to 2.9 percent in 2019 (although the small sample size should be noted; see Table 1). Most notably, the proportion of detainees who attributed their offending to methamphetamine increased significantly, from 6.2 percent in 2009 to 27.6 percent in 2019.

**Figure 1: Proportion of total sample who attributed offending to substance use, 2009 and 2019 (%)**




Note: Similar trends in attributions are observed when restricting the analyses to only detainees who had used alcohol and other drugs in the past month: with declines in alcohol (40.5% in 2009 to 29.1% in 2019) and heroin (54.3% in 2009 to 39.0% in 2019), increases in methamphetamine (33.1% in 2009 to 57.8% in 2019), and relatively no change for cannabis (13.5% in 2009 to 14.9% in 2019) and ecstasy (10.5% to 11.8%)

Source: AIC DUMA collection 2019 [computer file]; Payne & Gaffney (2012)

## Reasons for attributing offending to alcohol and illicit drugs

For detainees who used most drug types, the leading reason they attributed their offending to substance use was that they were high at the time of arrest (see Table 2). Almost two-thirds of methamphetamine-attributing detainees reported they were high (64.6%, *n*=388) compared with 18.6 percent (*n*=112) who reported they needed money to buy drugs. However, heroin-attributing detainees were equally likely to report that they had offended because they needed money to buy drugs (41.3%, *n*=26) as they were to report being high (41.3%, *n*=26). Attributing offending to the need for money to buy drugs suggests the offending relates to acquisitive crime, such as shoplifting or drug dealing. Detainees who said they were dependent on illicit drugs were much more likely than other detainees to attribute their offending to illicit drugs (58.3% vs 14.9%, $\chi^2(1)=465.7$, $p<0.001$).

Many of the 'other' reasons detainees cited related to the illegal nature of drug use, such as failing urinalysis testing and being arrested for possession and selling of the drug. To further examine whether the legality of drugs was the reason detainees attributed their offending to illicit drugs, we disaggregated data according to whether detainees were charged with a drug possession offence. The vast majority of detainees (85.1%) were not charged with a drug offence and, among those who were, most (87.8%) were also charged with another offence type. Detainees who were charged with a drug possession offence were more likely to attribute their offending to illicit drugs (65.7% of detainees for whom this was their only offence type, 65.2% of detainees also charged with other offence types) compared to detainees not charged with a possession offence (45.7%, $\chi^2(2)=36.2$, $p<0.001$). However, most detainees who attributed their offending to illicit drugs were not charged with drug-related offences (74.6%).

**Table 2: Reasons for attributing offending to drug use among detainees who had used in the past month**

| Substance type | I needed money to buy drugs | | I was high at the time | | I was hanging out for drugs | | Other | |
|---|---|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % | *n* | % |
| Cannabis | 17 | 11.4 | 81 | 54.4 | 17 | 11.4 | 45 | 30.2 |
| Heroin | 26 | 41.3 | 26 | 41.3 | 12 | 19.1 | 13 | 20.6 |
| Methamphetamine | 112 | 18.6 | 388 | 64.6 | 76 | 12.7 | 111 | 23.2 |
| Ecstasy | 0 | 0.0 | 4 | 40.0 | 0 | 0.0 | 4 | 80.0 |
| **Total** | **138** | **19.1** | **452** | **62.6** | **103** | **14.3** | **167** | **27.1** |

Note: Totals are calculated from the sample of detainees who had used that drug in the past month and who had attributed their offending to their drug use. Responses across multiple drug types were permitted

Source: AIC DUMA collection 2019 [computer file]

## Attributions by charge type

The proportion of crime attributable to drug and alcohol use was examined by charge type based on each detainee's principal or most serious offence (see Table 3). The subsample sizes for some crime types are small and these findings should be interpreted with caution. Similarly, we excluded ecstasy and heroin from Table 3 because too few detainees attributed their offending to these substances.

Drug or alcohol use contributed to offending among most detainees charged with homicide and related offences (66.7%); dangerous or negligent acts endangering persons, such as reckless driving (69.6%); robbery, extortion and related offences (58.9%); and unlawful entry with intent, breaking and entering, and related charges (59.5%). Some charge types were more likely to be attributed to illicit drugs than to alcohol and vice versa. More than half (57.8%) of charges relating to unlawful entry with intent were attributed to illicit drug use, compared to five percent attributed to alcohol use. Conversely, almost four times as many detainees charged with sexual assault or other sexual offences attributed their offending to alcohol (18.6%) than to illicit drugs (5.1%).

**Table 3: Principal offences attributed to alcohol and drugs, 2019 (%)**

| Type of offence | *n* | Alcohol | Meth-amphetamine | Cannabis | Illicit drugs | Any substance |
|---|---|---|---|---|---|---|
| Homicide and related offences | 12 | 33.3 | 41.7 | 0.0 | 41.7 | 66.7 |
| Acts intended to cause injury | 500 | 23.4 | 21.6 | 9.8 | 29.0 | 44.8 |
| Sexual assault and related offences | 59 | 18.6 | 5.1 | 0.0 | 5.1 | 23.7 |
| Dangerous or negligent acts endangering persons | 46 | 19.6 | 54.4 | 13.0 | 58.7 | 69.6 |
| Abduction, harassment and other offences against the person | 8 | 37.5 | 37.5 | 0.0 | 37.5 | 50.0 |
| Robbery, extortion and related offences | 56 | 37.5 | 16.1 | 7.1 | 25.0 | 58.9 |
| Unlawful entry with intent/burglary, break and enter | 116 | 5.2 | 48.3 | 9.5 | 57.8 | 59.5 |
| Theft and related offences | 443 | 9.7 | 36.8 | 6.6 | 42.4 | 49.2 |
| Fraud, deception and related offences | 43 | 7.0 | 18.6 | 0.0 | 18.6 | 23.3 |
| Illicit drug offences | 172 | 5.3 | 40.7 | 11.1 | 51.2 | 53.5 |
| Weapons and explosives offences | 35 | 8.6 | 14.3 | 8.6 | 22.9 | 31.4 |
| Property damage and environmental pollution | 41 | 24.4 | 7.3 | 4.9 | 12.2 | 34.2 |
| Public order offences | 98 | 34.7 | 15.3 | 4.1 | 18.4 | 45.9 |
| Traffic and vehicle regulatory offences | 106 | 19.8 | 12.3 | 2.8 | 14.2 | 30.2 |
| Offences against justice procedures, government security and government operations | 488 | 14.3 | 27.1 | 5.1 | 30.9 | 41.6 |
| **Total** | **2,306** | **16.4** | **27.8** | **7.0** | **33.5** | **45.4** |

Note: Totals exclude detainees with missing data on most serious offence (*n*=24). Miscellaneous offences (*n*=2) and charges attributed to MDMA (*n*=11) or heroin (*n*=64) are not reported separately due to low cell count

Source: AIC DUMA collection 2019 [computer file]

## Estimated number of principal offences attributable to substance use

The proportion of police detainees who attributed their offending to specific substances was used to estimate the approximate number of principal offences resulting in an offender being proceeded against that are attributable to drug or alcohol use. In 2018–19, the ABS recorded 345,286 principal offences allegedly committed by adult offenders. By applying the attributable fractions identified in the total sample of detainees, we estimate that alcohol use contributed to 56,627 principal offences, illicit drug use contributed to 115,671 principal offences, and all types of substance use contributed to 156,760 principal offences.

Attributable fractions were then applied to specific categories of crime (Table 4). For some categories, the number of offences could not be estimated due to small sample sizes or inconsistencies in the coding of specific offences between the DUMA dataset and the ABS recorded crime statistics dataset (see Table 4 note). Nonetheless, many high-harm offences were attributed to drug or alcohol use, including an estimated 480 homicides or related offences, 1,729 sexual assaults or related offences, and 30,956 assault or other acts intended to cause injury. Methamphetamine (which had the highest attribution rate among all substances analysed) was estimated to have contributed to approximately 95,990 offences, including 300 homicides and related offences, 3,962 unlawful entries with intent, and 12,414 offences relating to theft.

It is important to note that the estimated numbers of principal offences attributable to alcohol and other drugs are only approximations. These figures should be considered broad estimates, especially since the attributable fractions calculated from the DUMA dataset were applied to the recorded crime statistics based on the overarching ANZSOC division (highest level of classification) rather than specific ANZSOC subdivisions or groups.

**Table 4: Estimated number of principal offences attributable to alcohol and other drugs in 2018–19 (*n*)**

| Type of principal offence | Recorded offenders | Principal offences attributable to: | | | |
|---|---|---|---|---|---|
| | | Alcohol | Meth-amphetamine | Illicit drugs | All substances |
| Homicide and related offences | 720 | 240 | 300 | 300 | 480 |
| Acts intended to cause injury | 69,099 | 16,169 | 14,925 | 20,039 | 30,956 |
| Sexual assault and related offences | 7,263 | 1,351 | 370 | 370 | 1,729 |
| Abduction, harassment and other offences against the person | 3,551 | 1,332 | 1,332 | 1,332 | 1,776 |
| Robbery, extortion and related offences | 2,230 | 836 | 359 | 558 | 1,313 |
| Unlawful entry with intent/ burglary, break and enter | 8,202 | 427 | 3,962 | 4,741 | 4,880 |
| Theft and related offences | 33,733 | 3,272 | 12,414 | 14,303 | 16,597 |
| Fraud, deception and related offences | 11,064 | 774 | 2,058 | 2,058 | 2,578 |
| Illicit drug offences | 71,849 | 3,736 | 29,243 | 36,787 | 38,439 |
| Weapons and explosives offences | 14,134 | 1,216 | 2,021 | 3,237 | 4,438 |
| Property damage and environmental pollution | 11,751 | 2,867 | 858 | 1,434 | 4,019 |
| Public order offences | 49,275 | 17,098 | 7,539 | 9,067 | 22,617 |

Note: Traffic and vehicle regulatory offences are not included in the ABS recorded crime dataset. In the ABS datasets, dangerous or negligent acts endangering persons do not include those charged with dangerous or negligent operation of a vehicle (subdivision 041). Many DUMA detainees with a most serious offence of this type (dangerous or negligent acts endangering persons) had been charged with 041 subdivision offences, and as such, this offence category has not been calculated in the table above. Similarly, offences against justice procedures and offences against government security and government operations exclude charges relating to breaches of bail conditions (code 1523) in the ABS datasets. As many DUMA detainees' most serious offence related to breaches of bail, this category of offences (offences against justice procedures, and offences against government security and government operations) has not been included in the table above. In addition, the category of miscellaneous offences has not been calculated as very few DUMA detainees reported such offences as their most serious offence (*n*=2). Estimates rounded to nearest whole number

Source: AIC DUMA collection 2019 [computer file]; ABS recorded crime dataset (ABS; 2024)

We can cautiously use these data to estimate the number of offences in subsequent years. As we have already shown, rates of drug–crime attributions for certain illicit drugs have changed over time. In particular, the role of methamphetamine in criminal offending has increased, largely due to the increased use of the drug over this period. According to the most recent release from the ACIC's (2024) National Wastewater Drug Monitoring Program (Report 23), after falling during the pandemic, the consumption of illicit drug types measured as part of this study were at similar levels in 2024 as they were in 2019. Assuming that rates of drug–crime attribution for illicit drugs have also remained relatively stable since 2019, we estimate that approximately 98,451 principal offences in 2023–24—the latest year for which national data on recorded offenders were available—were related to illicit drugs (ABS 2025). Alcohol use has declined over this period (Tscharke et al. 2024); however, rates of alcohol-related victimisation have remained stable (AIHW 2024). Assuming this is also true for the involvement of alcohol in offending, we estimate there may have been up to 48,197 principal offences related to alcohol in 2023–24. The number of offences was smaller in 2023–24 than in 2018–19 due a decline in the total number of principal offences for many crime types.

## Discussion

More than half of detainees who participated in the DUMA program during 2019 and who had recently used at least one substance, and 45 percent of all detainees, attributed their current offending to illicit drug or alcohol use. Attribution rates varied for specific drugs, from ecstasy (0.5%), heroin (2.9%) and cannabis (6.9%), up to alcohol (16.5%) and methamphetamine (27.6%). These fractions are consistent with those reported by Canadian offenders (*n*=29,138) between 2006 and 2016, who also reported low rates for cannabis (5%) and opioids such as heroin (5%), and higher rates for alcohol (17%) and stimulants such as methamphetamine (31%).

Since 2009, attribution rates for alcohol use have declined by 13 percentage points. Most strikingly, there has been a fourfold increase in the reported role of methamphetamine use in detainees' offending since 2009 (from 6% to 28%). In 2019, 58 percent of detainees who had used methamphetamine in the past month attributed their offending to the drug. This trend is likely due, in part, to the rising prevalence of methamphetamine use among police detainees since 2009 (Voce et al. 2021; Voce & Sullivan 2022a). Changes in the preferred form of methamphetamine from powder to the more potent crystalline methamphetamine ('ice') may also have contributed to the increased attribution rate. Australian evidence suggests that, compared to powdered forms, use of crystalline methamphetamine is associated with more frequent use and dependence (Degenhardt et al. 2017), which in turn is associated with a greater risk of criminal involvement (McKetin et al. 2020).

Attributable fractions derived from samples of police detainees can be used to estimate the number of criminal offences committed in Australia that are related to drugs or alcohol use. Methamphetamine use alone was estimated to contribute to over 95,000 offences, whereas alcohol use contributed to over 56,000 offences. In total, more than 156,760 offences were attributable to alcohol or one of the illicit drug types measured as part of the DUMA program. This is likely an underestimation, given our results are limited to principal offences, adult offenders and certain illicit drug types.

These figures reflect the substantial societal harms that illicit drugs inflict on the Australian community, and the major criminal justice costs that must be expended to investigate and prosecute crimes linked to illicit drug use (National Drug Research Institute 2020). For example, offences attributable to substance use included almost 31,000 acts intended to cause injury, such as assaults, and most of these were attributed to methamphetamine use. This high rate of violent offences reflects the established link between illicit drug use and violence (Duke et al. 2018) and, more specifically, the link between heavier methamphetamine use among dependent individuals in Australia and violent behaviour (McKetin et al. 2014), including domestic violence (Morgan & Gannoni 2020).

Importantly, the vast majority of detainees who attributed their offending to drug use said they did so because the drug influenced their decision-making with regards to their offending. Very few detainees said they attributed their offending to their drug use because they had been arrested for possessing or selling drugs. Most detainees who were surveyed as part of the DUMA program were not charged with drug possession offences, and the majority of principal offences attributed to illicit drug use were not drug selling or possession offences or justice offences. Rather, they were offences against an individual or business, particularly property offences.

Our results therefore demonstrate the importance of reducing alcohol and illicit drug use among individuals who are involved in or at risk of offending behaviour. While substance use is one among many complex and interrelated factors that can contribute to criminal offending, compelling evidence shows that demand reduction programs, including programs targeting people in contact with the criminal justice system, can reduce drug-related crime and the cost to victims (Voce & Sullivan 2022b). These programs need to be delivered alongside targeted supply reduction efforts that are also supported by evidence (Mazerolle, Eggins & Higginson 2020).

## References

*URLs correct as at July 2025*

Australian Bureau of Statistics (ABS) 2024. Recorded Crime – Offenders, 2022–23. Canberra: ABS. https://www.abs.gov.au/statistics/people/crime-and-justice/recorded-crime-offenders/2022-23

Australian Bureau of Statistics 2023. National Offence Index (NOI). Canberra: ABS. https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-offence-classification-anzsoc/2023/national-offence-index-noi

Australian Bureau of Statistics 2011. Australian and New Zealand Standard Offence Classification (ANZSOC). Canberra: ABS. https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-offence-classification-anzsoc/2011

Australian Criminal Intelligence Commission (ACIC) 2024. *National Wastewater Drug Monitoring Program: Report 23.* Canberra: ACIC. https://www.acic.gov.au/publications/national-wastewater-drug-monitoring-program-reports/report-23-national-wastewater-drug-monitoring-program

Australian Institute of Health and Welfare (AIHW) 2025a. Alcohol, tobacco & other drugs in Australia. https://www.aihw.gov.au/reports/alcohol/alcohol-tobacco-other-drugs-australia

Australian Institute of Health and Welfare 2025b. National Drug Strategy Household Survey 2022–2023. https://www.aihw.gov.au/reports/illicit-use-of-drugs/national-drug-strategy-household-survey

Australian Institute of Health and Welfare 2024. Alcohol-related harms and risks in the NDSHS. https://www.aihw.gov.au/reports/alcohol/alcohol-related-harms-and-risks

Bennett T, Holloway K & Farrington D 2008. The statistical association between drug misuse and crime: A meta-analysis. *Aggression and Violent Behavior* 13(2): 107–118. https://doi.org/10.1016/j.avb.2008.02.001

Brown R & Morgan A 2019. The opioid epidemic in North America: Implications for Australia. *Trends & issues in crime and criminal justice* no. 578. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti04060

Darke S 2010. The toxicology of homicide offenders and victims: A review. *Drug and Alcohol Review* 29: 202–215. https://doi.org/10.1111/j.1465-3362.2009.00099.x

Degenhardt L et al. 2017. Crystalline methamphetamine use and methamphetamine-related harms in Australia. *Drug and Alcohol Review* 36(2): 160–170. https://doi.org/10.1111/dar.12426

Doherty L, Sullivan T & Voce A 2021. *Impact of the COVID-19 pandemic on cannabis demand and supply in Australia*. Statistical Bulletin no. 33. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sb78252

Duke AA, Smith KMZ, Oberleitner LMS, Westphal A & McKee SA 2018. Alcohol, drugs, and violence: A meta-meta-analysis. *Psychology of Violence* 8(2): 238–249. https://doi.org/10.1037/vio0000106

Foran HM & O'Leary KD 2008. Alcohol and intimate partner violence: A meta-analytic review. *Clinical Psychology Review* 28(7): 1222–1234. https://doi.org/10.1016/j.cpr.2008.05.001

Håkansson A & Jesionowska V 2018. Associations between substance use and type of crime in prisoners with substance use problems: A focus on violence and fatal violence. *Substance Abuse and Rehabilitation* 15(9): 1–9. https://doi.org/10.2147/SAR.S143251

Mazerolle L, Eggins E & Higginson A 2020. Street-level drug law enforcement: An updated systematic review. *Trends & issues in crime and criminal justice* no. 599. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti04640

McKetin R et al. 2020. The contribution of methamphetamine use to crime: Evidence from Australian longitudinal data. *Drug and Alcohol Dependence* 216: 108262. https://doi.org/10.1016/j.drugalcdep.2020.108262

McKetin R et al. 2014. Does methamphetamine use increase violent behaviour? Evidence from a prospective longitudinal study. *Addiction* 109(5): 798–806. https://doi.org/10.1111/add.12474

Morgan A & Gannoni A 2020. Methamphetamine dependence and domestic violence among police detainees. *Trends & issues in crime and criminal justice* no. 588. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti04190

Morgan A, Lyneham S, Davy D & Coughlan M 2018. *Towards national measures of alcohol-related crime*. Research Report no. 12. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/rr199916

Nagar M & Rabinovitz S 2024. Revisiting the alcohol-aggression link: The impact of alcohol consumption patterns. *Drug and Alcohol Dependence* 265: 112496. https://doi.org/10.1016/j.drugalcdep.2024.112496

National Drug Intelligence Center (NDIC) 2011. *The economic impact of illicit drug use on American society*. Washington, DC: United Sates Department of Justice. http://www.justice.gov/archive/ndic/pubs44/44731/44731p.pdf

National Drug Research Institute 2020. *Quantifying the social costs of cannabis use to Australia in 2015/16*. Perth: National Drug Research Institute, Curtin University. https://ndri.curtin.edu.au/publications-resources/project-reports-and-bulletins/social-and-economic-costs-of-substance-use

Parliamentary Joint Committee on Law Enforcement 2023. *Australia's illicit drug problem: Challenges and opportunities for law enforcement*. Canberra: Parliament of Australia. https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Law_Enforcement/IllicitDrugs

Payne J & Gaffney A 2012. How much crime is drug or alcohol related? Self-reported attributions of police detainees. *Trends & issues in crime and criminal justice* no. 439. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti251968

Payne J, Morgan A & Piquero A 2022. COVID-19 and social distancing measures in Queensland, Australia, are associated with short-term decreases in recorded violent crime. *Journal of Experimental Criminology* 18: 89–113. https://doi.org/10.1007/s11292-020-09441-y

Sutherland R et al. 2023. *Australian drug trends 2023: Key findings from the National Illicit Drug Reporting System (IDRS) interviews*. Sydney: National Drug and Alcohol Research Centre, UNSW Sydney. https://www.unsw.edu.au/research/ndarc/resources/australian-drug-trends-2023-key-findings-from-irds

Tscharke B et al. 2024. Seven-years of alcohol consumption in Australia by wastewater analysis: Exploring patterns by remoteness and socioeconomic factors. *Drug and Alcohol Dependence* 259: 111317. https://doi.org/10.1016/j.drugalcdep.2024.111317

Voce A & Sullivan T 2022a. *Drug use monitoring in Australia: Drug use among police detainees, 2021*. Statistical Report no. 40. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sr78573

Voce A & Sullivan T 2022b. What are the monetary returns of investing in programs that reduce demand for illicit drugs? *Trends & issues in crime and criminal justice* no. 657. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78733

Voce A, Sullivan T & Doherty L 2021. *Declines in methamphetamine supply and demand in Australia during the COVID-19 pandemic*. Statistical Bulletin no. 32. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sb78245

Voce I, Voce A, Morgan A & Sullivan T 2021. *The DUMA Drug Market Indicator Framework: Methamphetamine*. Statistical Report no. 33. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sr04879

Young M et al. 2021. Attributable fractions for substance use in relation to crime. *Addiction* 116: 3198–3205. https://doi.org/10.1111/add.15494

Zhong S, Yu R & Fazel S 2020. Drug use disorders and violence: Associations with individual drug categories. *Epidemiologic Reviews* 42(1): 103–116. https://doi.org/10.1093/epirev/mxaa006

This work was originally published as *Trends & issues in crime and criminal justice* no. 723.
It is available online at **https://doi.org/10.52922/ti78090**.

# 4. An experimental study of support for protest causes and tactics and the influence of conspiratorial beliefs

Anthony Morgan, Timothy Cubitt, Alexandra Voce and Isabella Voce

In the last few years protest activity has increased globally (van der Zwet et al. 2022). Australia has been no exception. Recent protest activity has advocated for a range of political and social causes, including support for increased action on climate change, opposition to public health measures introduced during the COVID-19 pandemic, women's safety, Indigenous deaths in custody and other more extreme ideological movements.

There were protests in response to lockdowns and social distancing measures from very early in the pandemic (McGowan 2020), including the emergence of the freedom movement in late 2020 (ABC News 2020). There was further protest activity following the rollout of the national vaccination program in early to mid-2021 (Noble 2021). While many of these protests were peaceful, there were numerous examples of disruptive behaviour and clashes between protesters and police, resulting in large numbers of fines and arrests, as well as the hospitalisation of police officers (Seyfort & Zagon 2021). This included notable anti-vaccine mandate protests in Melbourne, which attracted widespread condemnation (Marin-Guzman 2021).

As frustration with public health measures grew, antisocial behaviour and threats of violence towards political leaders escalated (Butler 2021). In early 2022, several thousand Australians travelled in a 'Convoy to Canberra' to protest against a range of grievances, organised using social media and supported through crowdfunding (Britton 2022). While fringe and conspiratorial views were apparent in some of the earliest protest activity, the convoy exemplified the central role that conspiratorial beliefs, the language of the sovereign citizen movement and other ideological motives had come to play in pandemic protest narratives (Khalil & Roose 2023; Roose 2022).

While COVID-19 protests garnered significant attention, protest activity occurred in support of several other movements in this same period. It too involved a mix of tactics. Following intense bushfires in the 2019–20 summer, tens of thousands of people protested nationally in support of action on climate change and environmental policy (Regan & Yeung 2020). Conversely, in mid-2022, Blockade Australia disrupted traffic heading into the centre of Sydney during peak hour, while concurrently protesting through Sydney's central business district (Thompson 2022). Despite involving no more than 60 individuals, there was significant disruption to the community and a number of protesters were arrested (McGuire 2022). Two large, nationally coordinated protests were also noteworthy. Marches against Indigenous deaths in custody and systemic racism occurred in mid-2020 (Henriques-Gomes & Visontay 2020), while the 'March 4 Justice' took place in March 2021 in 40 cities in response to high-profile cases of sexual violence, with estimates of up to 110,000 attendees (Gorman 2021). These were peaceful protests that, unlike much of the pandemic-related protest activity, were planned in consultation with authorities and held in accordance with public health measures.

In attempting to draw public attention to their social or political cause, protesters may use disruptive and at times aggressive tactics (Andrews & Caren 2010). Often this depends on the perceived legitimacy of the government of the day—protest violence is more likely when the government is perceived as coercive rather than legitimate (Sullivan 2018). Some protest groups can be unpredictable and volatile, attracting individuals with a greater propensity for violence (Maguire et al. 2016; Tyler et al. 2018). However, it is also clear that protest violence can be specific to the situation and is more likely to occur when the action is disorganised or spontaneous (Gustafson 2020; Ives & Lewis 2020), when there are counter-protesters present (Australian Security Intelligence Organisation 2020), and during periods of heightened emotion and tension following some sort of triggering event (Nassauer 2016). The use of more assertive policing measures can lead protesters to escalate to violence in competition with police tactics (della Porta 2014). But this is not always straightforward. What constitutes lawful or violent protest can vary between places and also over time, particularly in political and legal environments where tolerance for public protest may be lowered. This has implications for law enforcement, whose role it is to maintain civil order and enforce protest laws, and can shape how they respond but also how that response may be perceived.

The majority of Australians support the right to protest (Amnesty International 2022). However, public support for activism may be influenced by the tactics used by protesters, even among individuals who are personally aligned with the social or political cause (Feinberg, Willer & Kovacheff 2020; Muñoz & Anduiza 2019; Simpson, Willer & Feinberg 2018). Of course, what constitutes violent action is subjective, and can depend on the political beliefs of the individual (Hsiao and Radnitz 2020). Even so, the negative effects of violent protest have been shown in survey experiments (Simpson, Willer & Feinberg 2018) and in studies of real-world protest violence (Huet-Vaughn 2013; Muñoz & Anduiza 2019). Further, regardless of an individual's core ideology, violent protest action can reduce support for the cause overall, and diminish identification with a movement, with observers reporting the action to be immoral (Feinberg, Willer & Kovacheff 2020). This can lead to fewer protesters being willing to engage in subsequent protest activity—although violence by the state can have the opposite effect (Steinert-Threlkeld, Joo & Chan 2022).

Protest tactics can also be influenced by a group's underlying ideology. The onset of the pandemic saw an increase in protests endorsing conspiratorial views (Khalil & Roose 2023), which have been associated with increased support for violence (Uscinski et al. 2022). Recent research has found that individuals who show greater commitment to conspiratorial beliefs are also more likely to endorse ideological violence (Vegetti & Littvay 2021). Further, these beliefs are associated with support for political violence, particularly where they co-occur with mental health issues (Baum et al. 2023).

It is not clear how these findings apply to the Australian context, particularly as it relates to contemporary protest movements and the recent growth of certain ideological movements and conspiratorial beliefs. To measure differences in support for different protest causes and tactics, we conducted a randomised survey experiment. We examined scenarios involving fringe movements, including sovereign citizen and anti-lockdown and anti-vaccination protests (Khalil & Roose 2023), and more widely-accepted mainstream movements, including action on climate change (Bradley et al. 2022). We also examined the extent to which support was influenced by belief in conspiracy theories. This research has implications for both protest organisers seeking to generate community support for their movement and authorities who are responsible for ensuring that protest activity is lawful and non-violent.

# Method

## Sample

Data for this study were collected as part of a large national survey of online Australians about their political and social beliefs. Fieldwork was conducted by Roy Morgan Research in November and December 2022. Respondents were recruited from Roy Morgan's Single Source panel, which comprises individuals recruited through a rigorous clustersampled, face-to-face survey approach, and a panel managed by Pureprofile. The survey was sent to members of these panels aged 18 years and over who had joined to receive incentives in exchange for completing surveys.

Proportional quota sampling was used to ensure the sample broadly reflected the spread of people living in Australia. Quotas were based on the Australian adult population stratified by age, sex and usual place of residence, based on Australian Bureau of Statistics (2023) population data. The majority of respondents (58.7%) were drawn from Roy Morgan's Single Source survey panel. The raw completion rate for invitations sent to this panel was 7.1 percent; however, 78.3 percent of respondents who opened the invitation and were eligible to participate in the research went on to complete the survey. Poor-quality responses and duplicates were removed from the final sample.

The survey took respondents an average of 23.8 minutes to complete. The initial sample size was 13,302 respondents. One respondent was removed as they were present in more than one experimental group, leaving a sample of 13,301. Because we were interested in the differences in protest support between groups, and the randomisation ensured there were no threats to the internal validity of our findings, we did not apply post-stratification weights to the data.

## Experimental design

As part of the survey, respondents were asked to read descriptions of four protest scenarios:

- a group of pro-choice advocates marching to show support for protecting abortion access in Australia;
- a group of anti-vaccine, anti-lockdown and anti-mandate advocates protesting government-imposed mandatory COVID-19 vaccinations and lockdowns;
- a group of sovereign citizens protesting oppression by the government, which they view as an illegitimate corporation depriving them of their sovereign rights and imposing laws to which they do not consent; and
- a group of environmental activists protesting government inaction on climate change and calling for a new, green approach to energy and emissions.

A split-ballot design was embedded into the survey, in which respondents were randomly assigned to one of three mutually exclusive groups (Table 1). Each group was asked about different tactics used by protesters in support of each issue or movement. These included peacefully protesting (peaceful protest condition), disrupting peak-hour lunchtime traffic (disruptive protest condition), or violently clashing with police (violent protest condition). We did not vary the tactics used in the pro-choice protest scenario, which acted as a control condition to verify that the randomisation had been successful. Access to abortion services was chosen because it is an issue that has received consistently high support from the Australian public in the previous decade (IPSOS 2022). It is noteworthy that the survey was conducted in the months after the US Supreme Court overturned *Roe v Wade*, a landmark case that had legalised abortion nationwide (ABC News 2022), leading to an increased focus on the issue in Australia (Truu 2022), where most legal barriers (but not non-legal barriers) to abortion have been removed (Sifris & Penovic 2021).

The order in which scenarios were presented to each respondent was also randomised. This was done to avoid any carryover ordering effects, in which the respondents' ratings of support for or opposition to a given scenario (eg peaceful protests) could be influenced by the scenario that preceded it (eg violent protests). Respondents were asked to rate their level of support for each protest scenario on a scale from one (strongly oppose) to five (strongly support). Respondents could also answer that they did not know or decline to answer (combined into an unknown category).

**Table 1: Assignment of protest causes and tactics to experimental groups (*n*=13,301)**

| | Group A (*n*=4,463, 33.6%) | Group B (*n*=4,411, 33.2%) | Group C (*n*=4,427, 33.3%) |
|---|---|---|---|
| Pro-choice protest (control condition) | Marching | Marching | Marching |
| Environmental protest | Violently clash with police | Disrupt traffic | Peacefully protest |
| Sovereign citizen protest | Disrupt traffic | Peacefully protest | Violently clash with police |
| Anti-lockdown and anti-vaccination protest | Peacefully protest | Violently clash with police | Disrupt traffic |

## Other variables of interest

In addition to the control protest condition, respondents were asked to rate their level of agreement with the statement 'It's OK to support groups that use violence to fight injustices.' This was used to measure any differences between groups in their support for violence as a means of achieving political or social change.

To analyse the relationship between conspiratorial beliefs and support for different protest causes and tactics, respondents were asked the extent to which they agreed or disagreed with common conspiracies, some of which were related to the COVID-19 pandemic and public health measures. There were seven statements in total, ranging from more moderate beliefs to more extreme conspiracies:

- mainstream media reports in a biased and false manner;
- COVID-19 vaccine mandates are an unnecessary form of government control;
- COVID-19 has been intentionally released as part of an orchestrated plan;
- secret societies control the world;
- big events (wars, recessions, elections) are controlled by small groups secretly working against the rest of us;
- vaccines can be used maliciously to infect people with poison or insert microchips into people; and
- elites use microchips to track and affect the population's behaviour.

Responses were based on a five-point Likert scale ranging from strongly disagree to strongly agree. To produce a global measure of conspiracy beliefs, these ordinal scales were converted into a continuous variable (Sullivan & Artino 2013). These items had very good internal consistency ($\alpha$=0.87) and an overall mean score was calculated.

Finally, respondents were asked to provide a range of sociodemographic information as part of the survey. We also asked respondents whether they had ever participated in a protest or rally in-person in support of a political or social movement and, if so, whether they had participated in an in-person protest since January 2020. Respondents who said they had participated in an in-person protest since January 2000 were then asked the issues or movements they had supported through protest.

## Analytical approach

Our analysis proceeded in three stages. First, we examined whether the randomisation had resulted in a balanced sample in each of the three groups, and whether respondents in each of those three groups provided similar responses to common questions about protests and the use of violence. Second, we analysed the results from the experimental study, comparing the responses from the three groups to determine the extent to which support for protest activity varied according to the tactics employed. In both stages we used chi-square ($\chi^2$) tests of association to assess whether there was a statistically significant relationship between the experimental group assigned and the variable of interest. Where this result was statistically significant, we used Cramér's *V* to measure the strength of this association.

The third stage of the analysis examined whether the relationship between protest tactics and support for protests varied depending on whether a respondent believed in conspiracy theories. We did this by estimating a series of ordinary least square regression models. We converted the outcome variable—support for protests—from a five-point ordinal scale to a numeric variable (excluding respondents who did not know or declined to answer from this stage of the analysis). We included protest tactics as a categorical independent variable, with peaceful protest used as the reference category. We then included a variable that measured the strength of each respondent's beliefs in common conspiracy theories. In addition to estimating the main effects of these variables on the level of support for protests, we included an interaction between protest tactics and belief in conspiracy theories. These models also included several control variables that were expected to be related to support for protests, including age, gender, employment status, level of education, level of socio-economic disadvantage, relationship status, parental status, and whether the respondent had participated in protest activity (never protested, protested before January 2020 but not recently, and protested since January 2020).

## Limitations

The experimental design allowed us to understand how the level of support changes depending on whether a protest is peaceful, disruptive or violent. It ensures a high degree of internal validity. However, the findings cannot explain why a certain cause or tactic may garner more or less support. Moreover, although the protest scenarios described in the experiment were based on actual events in Australia, in which protesters halted peak-hour traffic (Thompson 2022) and violently clashed with police (Seyfort & Zagon 2021), it is possible that the hypothetical scenarios lack real-world context, which may have influenced the extent to which respondents indicated support (or not) for different protest causes and tactics. For example, support for protests that involve violent clashes with police may depend on the severity of that violence, how it started and how police responded, among other variables. Respondents may also have had limited awareness of the sovereign citizen movement beyond the short description provided in the survey. Finally, the current study was not intended to measure the general level of support in the community for protest activity, nor do the results apply to other social movements; rather, they are specific to the scenarios described in the survey. There are limitations associated with non-probability sampling, including not being able to generalise beyond the sample of respondents.

# Results

## Validating the randomisation

### *Sample characteristics*

Sample characteristics, presented in Table 2, show that the three experimental groups were very similar. There were no statistically significant differences between the groups in respondent age, Indigenous status, sexuality, place of birth, language spoken at home, disability, relationship status, parental status, usual place of residence, highest level of education, employment status or level of socio-economic disadvantage. The only statistically significance between the three groups was in the gender of respondents ($\chi^2(4)=10.6$, $p<0.05$, Cramér's $V=0.02$). Inspection of the adjusted residuals made it clear this was due to differences in the small proportion of respondents who identified as non-binary. Specifically, there was a lower than expected number of non-binary or other gender respondents in Group A, and a higher than expected number in Group C. When these respondents were excluded, the difference between groups in the proportion of respondents who were male or female was not statistically significant ($\chi^2(4)=3.1$, $p=0.21$). Ultimately, the difference between the groups was very small (demonstrated by a weak association), meaning it was unlikely to have had any effect on the results from the analysis that followed.

| Table 2: Sample characteristics, by experimental group (%) | | | | |
|---|---|---|---|---|
| | Group A | Group B | Group C | $\chi^2$, *p*-value |
| **Age** | | | | |
| 18–34 | 25.2 | 24.8 | 25.3 | 1.0, *p*=0.90 |
| 35–64 | 49.1 | 50.0 | 49.7 | |
| 65+ | 25.7 | 25.2 | 25.1 | |
| **Gender** | | | | |
| Male | 47.6 | 46.7 | 45.6 | 10.6, *p*<0.05 |
| Female | 52.2 | 52.8 | 53.8 | |
| Non-binary or other gender | 0.3 | 0.5 | 0.6 | |
| **First Nations** | 2.4 | 2.2 | 2.7 | 2.9, *p*=0.58 |
| **LGB+ respondents**[a] | 7.8 | 8.1 | 8.6 | 3.2, *p*=0.53 |
| **Born outside of Australia** | 23.3 | 23.5 | 22.8 | 2.7, *p*=0.60 |
| **Speaks a language other than English most often at home** | 5.6 | 4.9 | 5.6 | 5.1, *p*=0.28 |
| **Restrictive long-term health condition** | 12.7 | 13.2 | 12.0 | 3.1, *p*=0.54 |
| **Currently in a relationship** | 63.1 | 65.3 | 64.5 | 8.1, *p*=0.09 |
| **Children living at home** | 35.2 | 36.5 | 36.2 | 4.9, *p*=0.30 |
| **Usual place of residence (remoteness)** | | | | |
| Major city | 72.6 | 73.9 | 73.1 | 11.0, *p*=0.09 |
| Regional | 24.7 | 22.7 | 23.6 | |
| Remote | 2.0 | 2.8 | 2.6 | |
| **Highest level of education** | | | | |
| High school | 26.1 | 25.1 | 25.4 | 3.7, *p*=0.88 |
| Vocational | 22.8 | 23.4 | 23.3 | |
| University | 50.4 | 50.9 | 50.8 | |
| **Employment status** | | | | |
| Employed | 62.2 | 62.0 | 62.6 | 0.7, *p*=0.99 |
| Other | 33.5 | 33.9 | 33.1 | |
| Unemployed | 3.8 | 3.7 | 3.9 | |
| **Index of relative socio-economic disadvantage** | | | | |
| Quartile 1 (most disadvantaged) | 17.3 | 15.6 | 15.6 | 9.8, *p*=0.28 |
| Quartile 2 | 19.5 | 19.4 | 20.1 | |
| Quartile 3 | 25.3 | 25.5 | 26.2 | |
| Quartile 4 (least disadvantaged) | 37.1 | 38.9 | 37.4 | |

a: LGB+ respondents include those who identified as lesbian, gay, bisexual or another non-heterosexual sexual orientation

Note: Denominators include respondents who did not know or declined to answer the question. Where present, this unknown group was included in the comparison between groups

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

As shown in Table 3, the rate of protest participation, including since January 2020, did not significantly differ between the three experimental groups ($\chi^2(6)=8.4$, $p=0.21$). We also compared the groups in terms of the prevalence of participation in recent protests in support of specific issues or movements relevant to the current study. While participation in these protests was relatively rare , there were no differences between the groups in whether respondents had participated in protests in support of environmental issues or climate change ($\chi^2(2)=1.9$, $p=0.39$), women's safety or rights ($\chi^2(2)=1.4$, $p=0.50$), government policies on COVID-19 ($\chi^2(2)=1.1$, $p=0.58$), or the rejection of government authority and legal documents (typically associated with the sovereign citizen movement; $\chi^2(2)=0.5$, $p=0.77$).

**Table 3: Protest participation, by experimental group (%)**

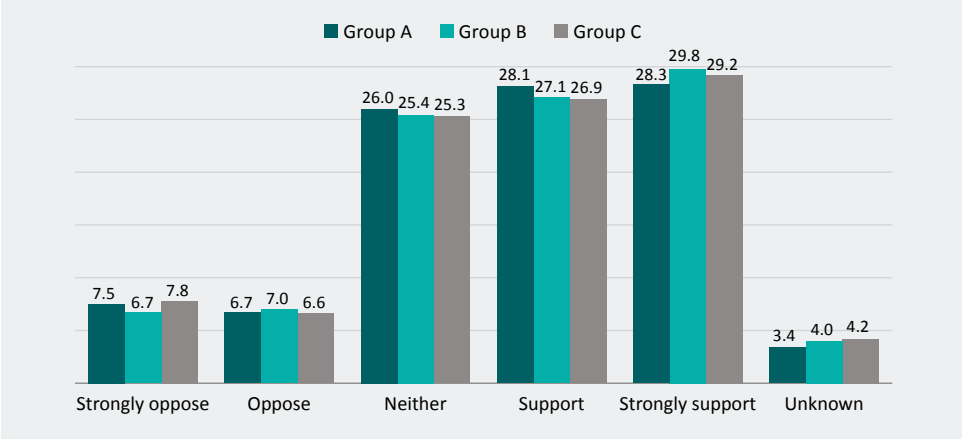
| | Group A | Group B | Group C | $\chi^2$, *p*-value |
|---|---|---|---|---|
| **Protest participation** | | | | |
| Never protested | 77.3 | 75.8 | 76.1 | 8.4, $p=0.21$ |
| Protested before January 2020, but not recently | 13.3 | 14.8 | 15.1 | |
| Protested since January 2020 | 7.0 | 7.2 | 6.7 | |
| Unknown | 2.4 | 2.2 | 2.2 | |
| **Participation in specific protest movements (since January 2020)** | | | | |
| Environment issues or climate change | 4.0 | 4.0 | 3.5 | 1.9, $p=0.39$ |
| Women's safety and rights | 2.1 | 2.4 | 2.3 | 1.4, $p=0.50$ |
| Government policies on COVID-19 | 1.3 | 1.5 | 1.3 | 1.1, $p=0.58$ |
| Rejection of government authority and legal documents | 0.4 | 0.5 | 0.4 | 0.5, $p=0.77$ |

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

### *Support for pro-choice protests and support for violence*

We then compared the groups on their support for pro-choice advocates marching to show support for protecting abortion access in Australia (Figure 1). There was no mention of specific tactics, and the same scenario was presented to all three groups. Support for this protest activity was high across all three groups, with only a small minority (13.7–14.4%) either opposed or strongly opposed. Importantly, there was no evidence that the level of support varied between the three experimental groups ($\chi^2(10)=11.4$, $p=0.33$).

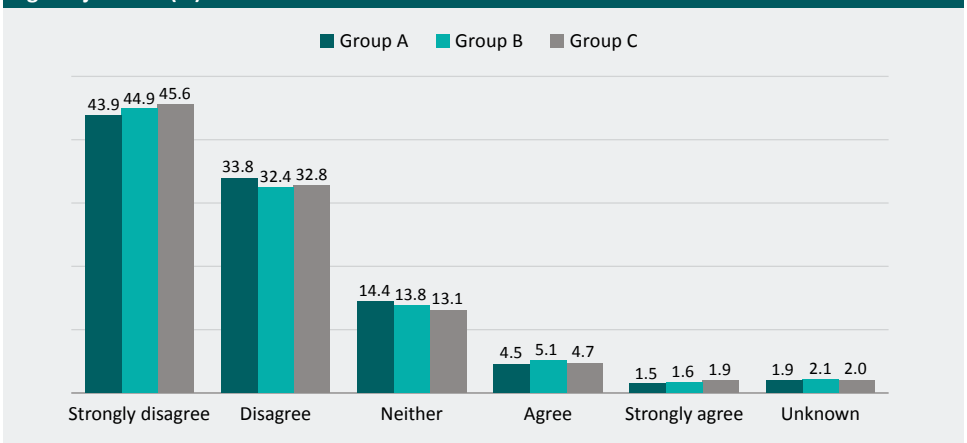
**Figure 1: Level of support for pro-choice protests (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

Next, we compared groups on their support for violence as a means of social or political change (Figure 2). Support for the use of violence to fight injustices was low among all three groups, with three-quarters of respondents (77.4–78.3%) either disagreeing or strongly disagreeing with the statement 'It's OK to support groups that use violence to fight injustices'. There were no significant differences in the level of support for this statement across the three groups ($\chi^2(10)=10.4$, $p=0.41$).

**Figure 2: Level of agreement with the statement 'It's OK to support groups that use violence to fight injustices' (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

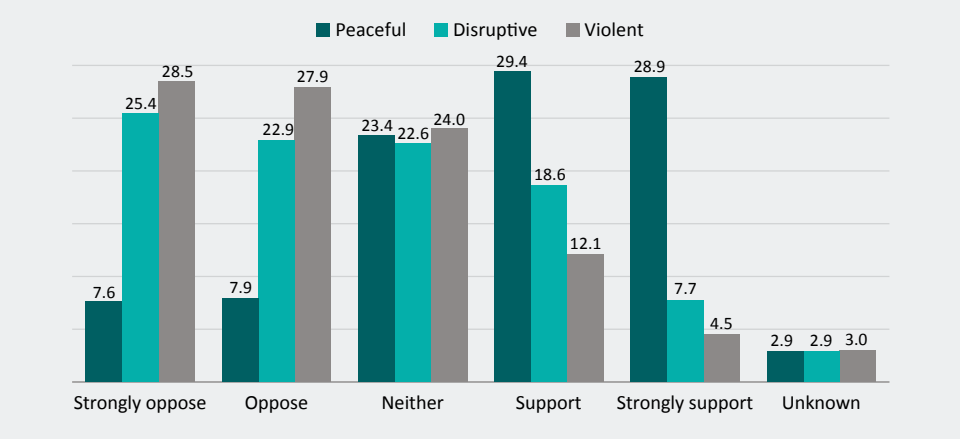
There were no major differences between the three groups in terms of sociodemographic characteristics and protest participation, and respondents in the three groups provided very similar responses to the same questions about pro-choice protests and the use of violence. This finding was important because it suggests that any observed differences in support between the different protest causes and tactics that follow are attributable to the experimental manipulation (ie varying the protest tactics). In other words, the randomisation was successful, and this was a valid experiment.

## Protest causes and tactics and community support

Among the three movements included, environmental protests had the highest level of support and the lowest level of opposition, irrespective of the tactics used. For example, more than half of respondents (58.3%) were supportive or strongly supportive of peaceful environmental protest, compared with 27.4 percent for sovereign citizen protests and 15.9 percent for anti-lockdown and anti-vaccination protests. The levels of support for and opposition to environmental protests were similar to those for pro-choice protests observed above.

There was clear evidence that protest tactics have a significant influence on support for protest activity. For example, in the case of environmental protests, only a small minority (15.5%) of respondents said they were opposed or strongly opposed to peaceful protests (Figure 3). In contrast, nearly half of respondents opposed or strongly opposed environmental protests that disrupted peak-hour traffic (48.3%), while more than half were opposed to environmental protests that involved violent clashes with police (56.4%). There was a very strong, statistically significant relationship between protest tactics and support for environmental protests ($\chi^2(10)=2{,}520.2$, $p<0.001$, Cramér's $V=0.31$).

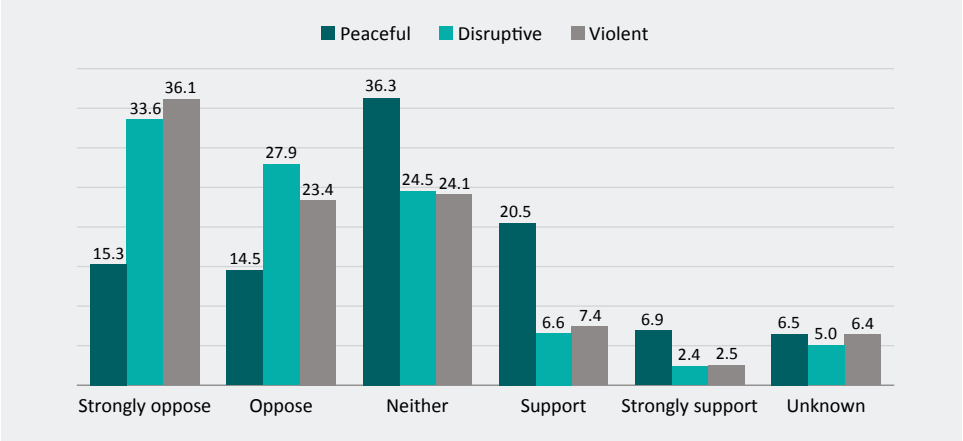
**Figure 3: Level of support for environmental protests, by protest tactic (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

Similarly, opposition to sovereign citizen protests was much higher when protests were described as disruptive (61.5%) or as involving violent clashes with police (59.5%) than when they were peaceful (29.7%). The levels of support for sovereign citizen protests were similar for disruptive protests and protests involving violent clashes with police. Nevertheless, there was a strong, statistically significant relationship between protest tactics and support for sovereign citizen protests ($\chi^2(10)=1{,}369.4$, $p<0.001$, Cramér's $V=0.23$; Figure 4).

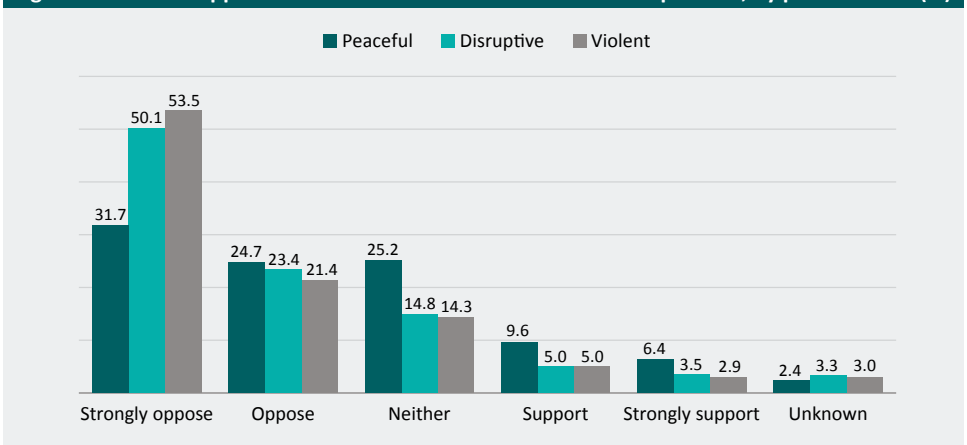
**Figure 4: Level of support for sovereign citizen protests, by protest tactic (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

Finally, while the level of support for anti-lockdown and anti-vaccination protests was lower than for other protest causes, there was still a relationship between protest tactics and support (Figure 5). While more than half the respondents were opposed to peaceful anti-lockdown and anti-vaccination protests (56.4%), this increased to three-quarters of respondents when protests were described as disruptive (73.5%) or as involving violent clashes with police (74.9%; $\chi^2(10)=646.9$, $p<0.001$, Cramér's $V=0.16$). Similar to sovereign citizen protests, there were almost identical levels of support for and opposition to disruptive protests and protests involving violent clashes with the police.

**Figure 5: Level of support for anti-lockdown and anti-vaccination protests, by protest tactic (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Belief in conspiracy theories and support for protest causes and tactics

The final stage of the analysis examined how belief in conspiracy theories influenced support for protest activity and, more specifically, whether the relationship between protest tactics and support for protests varied depending on whether a respondent believed in conspiracy theories. Results from three ordinary least square regression models are presented in Table 4. The statistically significant, negative coefficients for protest tactics show support was lower for disruptive and violent protest, relative to peaceful protest. The coefficient for conspiracy beliefs was significant in all models, but in different directions. A negative coefficient for the environmental protest model indicates that stronger conspiracy beliefs were associated with lower levels of support for environmental protest, while the positive coefficient for sovereign citizen and anti-lockdown and anti-vaccination protests indicates higher levels of support for protest among respondents with stronger conspiracy beliefs.

The interaction effect was significant in all three models. This was true for both coefficients in the environmental protest and sovereign citizen protest models, and the interaction between violent protest and conspiracy beliefs in the anti-lockdown and anti-vaccination protest model. These results show that not only does belief in conspiracy theories influence support for protests, independent of the tactics used, but it moderates the relationship between protest tactics and the level of support.

**Table 4: Ordinary least squares regression predicting level of support for protests**

| | Model 1: Environmental protests | | Model 2: Sovereign citizen protests | | Model 3: Anti-lockdown protests | |
|---|---|---|---|---|---|---|
| | *B* | *SE* | *B* | *SE* | *B* | *SE* |
| **Protest tactics (vs peaceful)** | | | | | | |
| Disruptive | −1.71*** | 0.07 | −0.99*** | 0.07 | −0.57*** | 0.06 |
| Violent | −2.25*** | 0.07 | −1.20*** | 0.07 | −0.68*** | 0.06 |
| **Conspiracy beliefs** | | | | | | |
| Strength of beliefs | −0.40*** | 0.02 | 0.31*** | 0.02 | 0.63*** | 0.02 |
| **Interactions** | | | | | | |
| Disruptive × conspiracy beliefs | 0.28*** | 0.03 | 0.09** | 0.03 | 0.04 | 0.03 |
| Violent × conspiracy beliefs | 0.41*** | 0.03 | 0.18*** | 0.03 | 0.06* | 0.03 |
| Constant | 4.78 | 0.07 | 2.41 | 0.07 | 0.76 | 0.04 |
| *n* | | 11,850 | | 11,525 | | 11,859 |
| *F* | | 367.33 | | 179.15 | | 266.3 |
| $R^2$ | | 0.31 | | 0.22 | | 0.33 |

***statistically significant at $p<0.001$, **statistically significant at $p<0.01$, *statistically significant at $p<0.05$

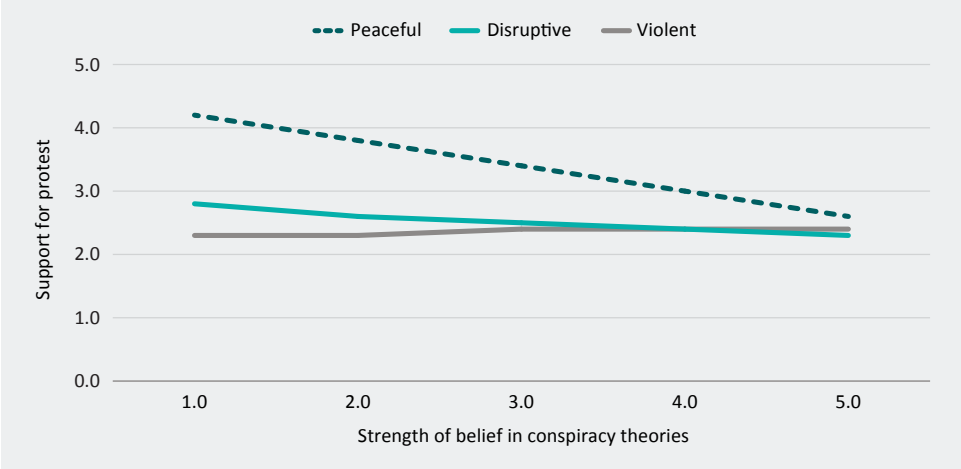
Note: Each model included controls for respondent age, gender, employment status, highest level of education, relationship status, whether children were living at home, and participation in protest. Robust standard errors were used

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

We can show how conspiracy beliefs influenced support for protests, and moderated the relationship between protest tactics and support for protests, by estimating the predictive margins. We did this for each model. These show the predicted level of support for protests at each level of belief in conspiracy theories, according to the type of protest tactic included in the scenario presented to the respondent.

Figure 6 shows that support for peaceful and disruptive environmental protests decreased as belief in conspiracy theories increased. Support for violent protest remained constant. Among respondents with the strongest belief in conspiracy theories, there was little difference in the levels of support for different protest tactics.

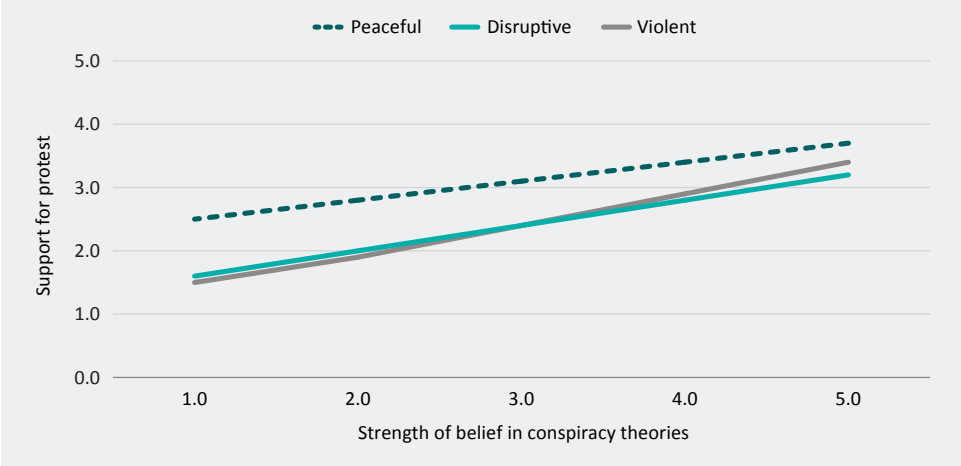
**Figure 6: Level of support for environmental protests, by belief in conspiracy theories and protest tactic (predictive margins)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

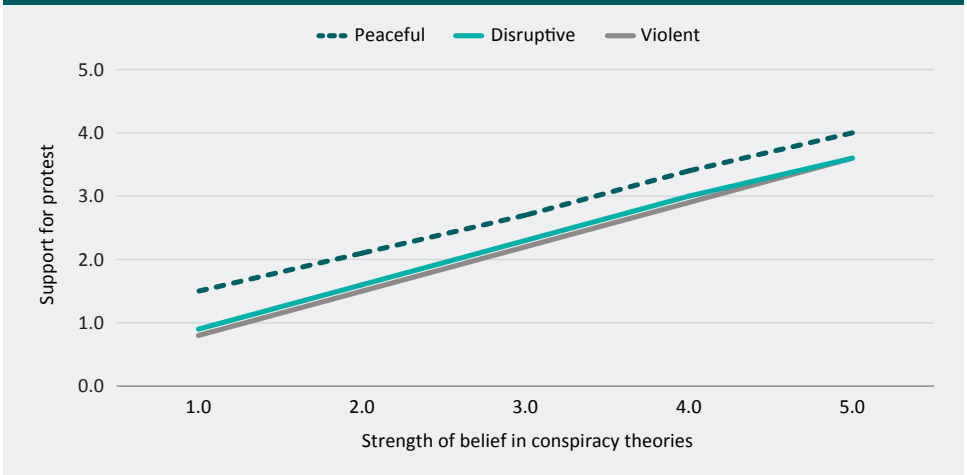
The effect of conspiracy beliefs on support for sovereign citizen protests and anti-lockdown and anti-vaccination protests is presented in Figure 7 and Figure 8, respectively. These show similar patterns. For both types of protest, respondents who held stronger beliefs in conspiracy theories were more likely to support protest activity. Further, support for violent protest increased at a faster rate than for peaceful protests (and, in the case of sovereign citizen protests, disruptive protest). This means conspiracy beliefs had a larger, positive effect on support for violent protest, relative to other protest tactics.

**Figure 7: Level of support for sovereign citizen protests, by belief in conspiracy theories and protest tactic (predictive margins)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

**Figure 8: Level of support for anti-lockdown and anti-vaccination protests, by belief in conspiracy theories and protest tactic (predictive margins)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Discussion

Our study shows that community support for protests varied according to the social issue or movement and the tactics used by protesters. The strongest opposition was to anti-lockdown and anti-vaccination protests, followed by protests relating to the sovereign citizen movement. Protests about environmental issues had the most support. This largely reflects popular opinion on these issues. Support for action on climate change is very high in the general population (Australian Conservation Foundation 2021). Similarly, there was widespread support for Commonwealth and state and territory public health measures during the COVID-19 pandemic and worry about the effects of the disease (Australian Bureau of Statistics 2021; Tranter 2022). Further, many anti-lockdown protests took place in contravention of public health orders (Khalil & Roose 2023). There were also numerous examples of disruptive behaviour and violence occurring during protests against pandemic-related public health measures (Butler 2021; Marin-Guzman 2021). There was extensive media coverage of disruptive environmental protest (eg Thompson 2022) immediately prior to our survey, but this involved a small group of activists. Similarly, the sovereign citizen movement is largely a fringe movement associated with strong anti-government views (Khalil & Roose 2023) and was prominent in a large-scale convoy protest in Canberra, alongside more malicious actors (Roose 2022), which was widely condemned by the local community. Once a protest movement is associated with unlawful, violent or disruptive activity, it may be that the loss of support among the general public endures, even when subsequent protest activity is peaceful.

Respondents were significantly more likely to oppose violent or disruptive protests than peaceful protests, regardless of the issue or movement being protested. There was generally very little difference in levels of support between disruptive protests and those involving violent clashes with police. These findings are consistent with evidence that protest tactics considered to be extreme (harmful, disruptive or violent) can be counterproductive, alienating both supporters and neutral observers and reducing broader community support for the cause (Feinberg, Willer & Kovacheff 2020; Simpson, Willer & Feinberg 2018). When protesters adopt such tactics, the public may perceive the movement as representing the extreme viewpoints of the group. In turn, this may lead to members of the public, even those who sympathise with a movement's goals, feeling disconnected from the movement—believing that the protesters do not represent them. Some activists may be unaware of this effect, believing instead that there is wider public support for their movement, irrespective of the tactics used (Feinberg, Willer & Kovacheff 2020).

Not all protesters attend protest activity with the same intention, particularly among larger protests. Protest activity may be unpredictable and, when a cause has been associated with disruptive or violent protest in the past, it may attract individuals who attend because of the possibility of violence (Maguire et al. 2016; Tyler et al. 2018). Early protest activity relating to COVID-19 regulations was comparatively peaceful and, while some activism focused on vaccination (Khalil & Roose 2023), there were also industrial and employment issues that attracted peaceful protest activity (Marin-Guzman 2021; Seyfort & Zagon 2021). However, with the onset of disruptive and violent protest in mid- to late 2021, it is possible that disruption and violence became associated with anti-lockdown activism. Our findings support the notion that more extreme protest tactics tend to undermine support, even among those who are aligned with the cause (Feinberg, Willer & Kovacheff 2020; Simpson, Willer & Feinberg 2018).

There are important lessons from this research for those who engage in activism to effect societal or policy changes. Not only are peaceful protests more successful at generating support in the community, they have also been shown to be more effective at promoting policy change than violent protests (Huet-Vaughn 2013). Disruptive or violent protest, regardless of the cause, is less likely to garner support from the community, while also risking recruiting individuals who are more interested in violent or disruptive behaviour than the cause itself (Maguire et al. 2016; Tyler et al. 2018). The effects depend on who is responsible for the violence and the scale of that violence (Steinert-Threlkeld, Joo & Chan 2019). The primary goal of some movements may not be to generate community support. Indeed, communicating a grievance with like-minded people is often the dominant motive, even if it is unlikely to lead to change (van Stekelenburg, Klandermans & van Dijk 2011). Nevertheless, the results encourage the use of peaceful protest where the aim is to shape policy. Whether these findings apply to other protest causes is unclear.

Separately, these findings may be important for engagement and messaging by authorities. In Australia the right to peaceful assembly is protected under international conventions and some state and territory legislation (Martin 2021). Communication with protest organisers and, where possible, attendees may help promote a shared understanding that peaceful protest activity is the best mechanism for generating wider community support for the cause and improving the chances of influencing public policy. Communication of this type may also offer an opportunity to identify and de-escalate extreme or fringe factions ahead of planned protest activity (Police Executive Research Forum 2022), which would benefit both authorities and protest organisers.

Lastly, our finding that belief in conspiracy theories significantly influences support for all forms of protest demonstrates the important role that underlying ideologies play in shaping a person's attitudes towards protest activity (Thomas et al. 2024). Belief in conspiracy theories has been shown to reduce support for policies to address climate change (Biddlestone, Azevedo & van der Linden 2022), intentions to engage in politics (Jolley & Douglas 2014) and extreme right-wing beliefs often associated with denial of climate change (Imhoff et al. 2022), which may explain why an increase in conspiratorial beliefs reduced support for peaceful environmental protests. That it increased support for disruptive and violent protest as part of more fringe movements, over and above the effect on support for peaceful protests, is also significant. This adds to the growing evidence of the links between conspiratorial beliefs and support for the use of violence in anti-government social movements (Thomas et al. 2024; Uscinski et al. 2022; Vegetti & Littvay 2021). It also supports prior research which found that conspiratorial beliefs are associated with greater willingness to justify violence and support or participate in violent or unlawful behaviour on behalf of one's own group (Morgan, Cubitt & Voce 2024; Rottweiler & Gill 2022). Given the proliferation of conspiracy theories that accompanied the pandemic (Waldek, Droogan & Ballsun-Stanton 2022), and persisted as pandemic regulations eased, there is a need to draw on the growing evidence on how to detect disinformation online (Aïmeur, Amri & Brassard 2023) and effectively communicate information about social or policy issues that are frequently targeted by malicious actors (Lazić & Žeželj 2021). This may help to reduce support for violent protest among those susceptible to this disinformation without infringing on people's right to protest.

## References

*URLs correct as at May 2024*

ABC News 2022. United States supreme court overturns Roe v Wade, erasing constitutional right to abortion. https://www.abc.net.au/news/2022-06-25/united-states-supreme-court-overturns-roe-v-wade/101183036

ABC News 2020. People protesting against coronavirus lockdown arrested in Melbourne amid clashes with police. https://www.abc.net.au/news/2020-09-13/victoria-coronavirus-anti-lockdown-protest-arrests-melbourne/12659410

Aïmeur E, Amri S & Brassard G 2023. Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining* 13(1). https://doi.org/10.1007/s13278-023-01028-5

Amnesty International 2022. Amnesty International Australia 2022 Human Rights Barometer. https://www.amnesty.org.au/report-amnesty-international-australia-human-rights-barometer-2022/

Andrews K & Caren N 2010. Making the news movement organizations, media attention, and the public agenda. *American Sociological Review* 75(6): 841–866. https://doi.org/10.1177/0003122410386689

Australian Bureau of Statistics 2023. National, state and territory population, September 2022. https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/latest-release

Australian Bureau of Statistics 2021. Household impacts of COVID-19 survey, June 2022. https://www.abs.gov.au/statistics/people/people-and-communities/household-impacts-covid-19-survey/jun-2021

Australian Conservation Foundation 2021. Australia's biggest climate poll. https://d3n8a8pro7vhmx.cloudfront.net/auscon/pages/19310/attachments/original/1630276451/PollingReport.pdf?1630276451

Baum MA, Druckman JN, Simonson MD, Lin J & Perlis RH 2023. The political consequences of depression: How conspiracy beliefs, participatory inclinations, and depression affect support for political violence. *American Journal of Political Science* 68(2): 575–594. https://doi.org/10.1111/ajps.12827

Biddlestone M, Azevedo F & van der Linden S 2022. Climate of conspiracy: A meta-analysis of the consequences of belief in conspiracy theories about climate change. *Current Opinion in Psychology* 46: 101390. https://doi.org/10.1016/j.copsyc.2022.101390

Bradley G, Deshpande S, Foxwell-Norton K, Hennessey N & Jackson M 2022. *Griffith climate action survey: Summary for policy and decision making*. Brisbane: Griffith University. https://www.griffith.edu.au/research/climate-action/national-longitudinal-survey

Britton A 2022. What you need to know about the 'freedom convoys' seen in Australia and across the world. *SBS News*, 16 February. https://web.archive.org/web/20220220223459/https://www.sbs.com.au/news/what-you-need-to-know-about-the-freedom-convoys-seen-in-australia-and-across-the-world/f992191c-8f41-43ae-979f-e0514d44ff7b

Butler J 2021. 'Sick and scary': Condemnation of gallows at anti-Andrews protests. https://thenewdaily.com.au/news/2021/11/16/gallows-daniel-andrews-protest/

della Porta D 2014. Competitive escalation during protest cycles: Comparing left-wing and religious conflicts. In C Demitriou (ed), *Dynamics of political violence: A process oriented perspective on radicalisation and the escalation of political conflict*. London: Routledge

Feinberg M, Willer R & Kovacheff C 2020. The activists dilemma: Extreme protest actions reduce popular support for social movements. *Journal of Personality and Social Psychology* 119(5): 1086–1111. https://doi.org/10.1037/pspi0000230

Gorman A 2021. 'Enough is enough!' Where, when and why March 4 Justice protests are taking place across Australia. https://www.theguardian.com/australia-news/2021/mar/15/march-4-justice-where-when-womens-march-protests-australia-march4justice-sydney-melbourne-canberra-brisbane

Gustafson D 2020. Hunger to violence: Explaining the violent escalation of nonviolent demonstrations. *Journal of Conflict Resolution* 64(6): 1121–1145. https://doi.org/10.1177/0022002719890669

Henriques-Gomes L & Visontay E 2020. Australian Black Lives Matter protests: Tens of thousands demand end to Indigenous deaths in custody. *The Guardian*, 6 June. https://www.theguardian.com/australia-news/2020/jun/06/australian-black-lives-matter-protests-tens-of-thousands-demand-end-to-indigenous-deaths-in-custody

Hsiao Y & Radnitz S 2020. Allies or agitators? How partisan identity shapes public opinion about violent or nonviolent protests. *Political Communications* 38(4): 479–497. https://doi.org/10.1080/10584609.2020.1793848

Huet-Vaughn E 2013. *Quiet riot: The causal effect of protest violence*. https://doi.org/10.2139/ssrn.2331520

Imhoff R et al. 2022. Conspiracy mentality and political orientation across 26 countries. *Nature Human Behaviour* 6: 392–403. https://doi.org/10.1038/s41562-021-01258-7

IPSOS (Institut Public de Sondage d'Opinion Secteur) 2022. *Global views on abortion survey 2022*. https://www.ipsos.com/en/global-advisor-abortion-2022

Ives B & Lewis JS 2020. From rallies to riots: Why some protests become violent. *Journal of Conflict Resolution* 64(5): 958–986. https://doi.org/10.1177/0022002719887491

Jolley D & Douglas KM 2014. The social consequences of conspiracism: Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint. *British Journal of Psychology* 105: 35–56. https://doi.org/10.1111/bjop.12018

Khalil L & Roose J 2023. Anti-government extremism in Australia: Understanding the Australian anti-lockdown freedom movement as a complex anti-government social movement. *Perspectives on Terrorism* 17(1): 144–169. https://doi.org/10.19165/DUYA5041

Lazić A & Žeželj I 2021. A systematic review of narrative interventions: Lessons for countering anti-vaccination conspiracy theories and misinformation. *Public Understanding of Science* 30(6): 644–670. https://doi.org/10.1177/09636625211011881

Maguire E, Barak M, Cross K & Lugo K 2016. Attitudes among Occupy DC participants about the use of violence against police. *Policing and Society* 28(5): 526–540. https://doi.org/10.1080/10439463.2016.1202247

Marin-Guzman D 2021. Victoria to close down construction industry after wild protests. *Australian Financial Review*, 20 September. https://www.afr.com/work-and-careers/workplace/anti-vaxxer-tradies-clash-with-cfmeu-officials-20210920-p58t7i

Martin G 2021. Protest, policing and law during COVID-19: On the legality of mass gatherings in a health crisis. *Alternative Law Journal* 46(4): 275–281. https://doi.org/10.1177/1037969X211029963

McGowan M 2020. Ten arrested and police officer injured at protest against Victoria's Covid-19 lockdown laws. *The Guardian*, 10 May. https://www.theguardian.com/australia-news/2020/may/10/ten-arrested-and-police-officer-injured-at-protest-against-victorias-covid-19-lockdown-laws

McGuire A 2022. Car drives into protest by climate activist group in Sydney CBD. *Sydney Morning Herald*, 27 June. https://www.smh.com.au/national/nsw/sydney-harbour-tunnel-closed-during-blockade-australia-protest-20220627-p5awsz.html

Morgan A, Cubitt T & Voce I 2024. *Participation in anti-authority protests and vulnerability to radicalisation*. Research Report no. 31. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/rr77413

Muñoz J & Anduiza E 2019. 'If a fight starts, watch the crowd': The effect of violence on popular support for social movements. *Journal of Peace Research* 56(4): 485–498. https://doi.org/10.1177/0022343318820575

Nassauer A 2016. From peaceful marches to violent clashes: A micro-situational analysis. *Social Movement Studies* 15(5): 515–530. https://doi.org/10.1080/14742837.2016.1150161

Noble F 2021. Anti-coronavirus vaccine protests held in Melbourne, Sydney, Brisbane, Adelaide, Perth. *9 News*, 21 February. https://www.9news.com.au/national/coronavirus-protest-melbourne-brisbane-sydney-pete-evans/a328629b-4d7f-48b1-a13f-72fe42dd8c2b

Police Executive Research Forum 2022. *Rethinking the police response to mass demonstrations: 9 recommendations*. Critical Issues in Policing Series. Washington, D.C.: Police Executive Research Forum. https://www.policeforum.org/critical-issues-series

Regan H & Yeung J 2020. Tens of thousands protest Australian PM's climate policies amid bushfire crisis. CNN, 10 January. https://edition.cnn.com/2020/01/10/australia/australia-fires-climate-protest-morrison-intl-hnk/index.html

Roose J 2022. How 'freedom rally' protesters and populist right-wing politics may play a role in the federal election. *The Conversation*, 15 February. https://theconversation.com/how-freedom-rally-protesters-and-populist-right-wing-politics-may-play-a-role-in-the-federal-election-176533

Rottweiler B & Gill P 2022. Conspiracy beliefs and violent extremist intentions: The contingent effects of self-efficacy, self-control and law related morality. *Terrorism and Political Violence* 34(7): 1485–1504. https://doi.org/10.1080/09546553.2020.1803288

Seyfort S & Zagon C 2021. More than 200 arrests made on third day of Melbourne protests. Nine News, 23 September. https://www.9news.com.au/national/victoria-protest-update-coronavirus-daniel-andrews-pleads-for-vaccinations-amid-construction-rallies/b6f176fa-3dd8-4d4a-bfcc-2ffd0da93b92

Sifris R & Penovic T 2021. Barriers to abortion access in Australia before and during the COVID-19 pandemic. *Women's Studies International Forum* 86: 102470. https://doi.org/10.1016/j.wsif.2021.102470

Simpson B, Willer R & Feinberg M 2018. Does violent protest backfire? Testing a theory of public reactions to activist violence. *Socius* 4. https://doi.org/10.1177/2378023118803189

Steinert-Threlkeld Z, Joo J & Chan A 2022. How state and protester violence affect protest dynamics. *Journal of Politics* 84(2): 798–813. https://doi.org/10.1086/715600

Sullivan H 2018. Sticks, stones, and broken bones: Protest violence and the state. *Journal of Conflict Resolution* 63(3): 700–726. https://doi.org/10.1177/0022002718763932

Sullivan G & Artino A 2013. Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education* 5(4): 541–542. https://doi.org/10.4300/JGME-5-4-18

Thomas EF et al. 2024. Do conspiracy beliefs fuel support for reactionary social movements? Effects of misbeliefs on actions to oppose lockdown and to "stop the steal". *British Journal of Social Psychology*. Advance online publication. https://doi.org/10.1111/bjso.12727

Thompson G 2022. Blockade Australia shut down Sydney with climate change protests. Now they're fighting arrests in court. *ABC News*, 14 August. https://www.abc.net.au/news/2022-08-14/blockade-australia-arrests-for-protesting-on-road-harbour-bridge/101320368

Tranter BK 2022 Your money or your life? Public support for health initiatives during the COVID-19 pandemic. *Australian Journal of Social Issues* 57: 544–561. https://doi.org/10.1002/ajs4.220

Truu M 2022. How Australia's version of Roe v Wade, a little known court case, set a national precedent for reproductive rights. *ABC News*, 29 June. https://www.abc.net.au/news/2022-06-29/australias-roe-v-wade-menhennitt-ruling-abortion-access/101185984

Tyler D, Barak M, Maguire E & Wells W 2018. The effects of procedural injustice on the use of violence against police by Occupy Wall Street protesters. *Police Practice and Research* 19(2): 138–152. https://doi.org/10.1080/15614263.2018.1418153

Uscinski J et al. 2022. The psychological and political correlates of conspiracy theory beliefs. *Scientific Reports* 12: 21672. https://doi.org/10.1038/s41598-022-25617-0

van der Zwet K, Barros A, van Engers T & Sloot P 2022. Emergence of protests during the COVID-19 pandemic: Quantitative models to explore the contributions of societal conditions. *Humanities and Social Sciences Communications* 9: 68. https://doi.org/10.1057/s41599-022-01082-y

Van Stekelenburg J, Klandermans B & van Dijk WW 2011. Combining motivations and emotion: The motivational dynamics of collective action participation. *International Journal of Social Psychology* 26(1): 91–104. https://doi.org/10.1174/021347411794078426

Vegetti F & Littvay L 2021. Belief in conspiracy theories and attitudes toward political violence. *Italian Political Science Review* 52(1): 18–32. https://doi.org/10.1017/ipo.2021.17

Waldek L, Droogan J & Ballsun-Stanton B 2022. *Online far right extremist and conspiratorial narratives during the COVID-19 pandemic*. https://doi.org/10.5281/ZENODO.5732611

This work was originally published as *Trends & issues in crime and criminal justice* no. 702. It is available online at **https://doi.org/10.52922/ti77611**.

# 5. Exposure to and sharing of fringe or radical content online

Timothy Cubitt and Anthony Morgan

Violent extremists use the internet to communicate with each other and with the general public, to distribute propaganda, radicalise and recruit at-risk individuals, raise funds and plan attacks (Broadhurst et al. 2017; Droogan & Waldek 2018). For example, both Islamist and right-wing extremists, including those involved in committing extremist violence, have been shown to be active online (Gaudette, Scrivens & Venkatesh 2022; Whittaker 2021). Research has focused on the supply side of online extremist content, including how extremist groups use the internet to disseminate radical content (Bastug, Douai & Akca 2020; Droogan, Waldek & Blackhall 2018). Many extremists have moved online because it gives them access to a larger audience and more potential recruits (Broadhurst et al. 2017; Droogan & Waldek 2018). They are selective in the online platforms they use, making choices based on cost, security, reach and user-friendliness (Walther & McCoy 2021). Some mainstream social media platforms are used primarily to reach large audiences, while other more niche platforms are used to circumvent content removal mechanisms and share information securely.

Exposure to and engagement with extremist actors and content online is a risk factor for radicalisation (Frischlich 2021; Reeve 2023). In their large-scale systematic review, Wolfowicz and colleagues (2021) found that exposure to radical content online was a risk factor for attitudes that support the use of violent radical behaviours, while having deviant or radical peers and online contact with extremists was associated with radical behaviours. However, exposure to radical content is usually regarded as a contributory factor to radicalisation rather than a direct cause (Gill et al. 2017), with research focusing on how exposure occurs and how online behaviours can lead to radicalisation (Wolbers et al. 2023). While the internet expands the reach and opportunity of those wanting to distribute radical content, we cannot yet be sure of the conditions under which individuals who hold radical attitudes, supported by online content, mobilise to violence. Recent evidence suggests that exposure to radical content alone does not necessarily lead to radicalisation or violence; however, the increased exposure afforded by the internet may ultimately result in a far larger pool of vulnerable individuals being exposed to radical content (Hassan et al. 2018; Nienierza et al. 2021). Despite this, most extremists are thought to initially encounter this content while searching for it rather than through chance encounters (Nienierza et al. 2021). The risk of radicalisation is considered greater when radical content is intentionally accessed than following passive exposure (Frissen 2021).

Most current research into this area has relied on data extracted from social media platforms (Droogan et al. 2022; Droogan, Waldek & Blackhall 2018). While useful in understanding the type of information shared and behaviours of users active on these platforms, these methods do not establish the extent to which online users are exposed to or intentionally seek out radical content. Similarly, little is known about the online activity of those who go on to share that content, join groups that promote extreme or radical views and invite others to join those groups. In this chapter we analyse the characteristics and online behaviours of Australians who access radical content online and the behaviours of those who share radical content and join radical groups who propagate that content.

# Methodology

## Sample

This research uses data from a large sample of online Australians surveyed about their political and social beliefs (*n*=13,302). Conducted by Roy Morgan Research in late 2022, the survey was sent to panels of individuals aged 18 years and over who had voluntarily joined to receive incentives in exchange for completing surveys. Participants were first recruited from Roy Morgan's Single Source survey panel, which is recruited through a rigorous cluster sampled, face-to-face survey approach. Proportional quota sampling was used to ensure the sample was broadly reflective of the spread of people living in Australia. Quotas were based on the Australian adult population stratified by age, sex and usual place of residence, based on population data from the Australian Bureau of Statistics (2023). Weighting by age, sex and usual place of residence was undertaken to ensure that the data were representative of the spread of the Australian population. Additional random iterative method weights were applied to correct for education level, internet and social media use.

Respondents were asked whether, in the last 12 months, they had seen political, ideological or religious content online that mainstream society or the government might describe as fringe, unorthodox or radical (hereafter referred to as 'fringe or radical content'). This included online messages, posts or discussions; videos they had watched, shared on bulletin boards or attached to text or email messages (excluding fictional films); and books, magazines, articles or other written material, downloaded to print or received attached to a message such as an email. This latter category included fringe or radical content in the form of political or religious content such as speeches, sermons, manifestos, policies, ideas or information designed to influence respondents' views on world events towards an identifiable ideology or issue. Questions focused on the type of content, whether access had been intentional, and whether respondents had then shared the content with others. Additionally, we asked respondents whether they had ever accessed any of this type of content offline. Finally, respondents were asked whether they had encountered online groups, communities, servers, clubs or forums online (hereafter referred to as 'online groups') that mainstream society or the government might describe as fringe or radical and, if so, whether they had accessed them intentionally or joined the group.

Responses to these questions were used to develop the groups that are the focus of this research:

- respondents who had seen fringe or radical content online in the 12 months prior to the survey but had done so unintentionally (*n*=3,792);
- respondents who had intentionally accessed this content (*n*=1,253);
- respondents who had shared this content (whether they had accessed it intentionally or not, *n*=522); and
- respondents who had intentionally accessed online groups that mainstream society or the government might describe as fringe or radical (*n*=524).

Respondents were excluded from the analysis if they did not answer questions about whether they had viewed fringe or radical content, whether they had done so intentionally, or whether they had shared the content or joined fringe or radical groups.

## Analytic approach

The analysis proceeded in several stages. First, we explored the extent and types of content respondents had accessed, unintentionally or intentionally, and the sharing of this content online. We then examined the characteristics of respondents who accessed and shared fringe or radical content, how and when they were first exposed to this material, and their online activity. We then focused specifically on the platforms used to share fringe or radical content online, and the online fringe or radical groups respondents accessed or joined.

Respondents who had seen fringe or radical content unintentionally were compared with respondents who had not seen it to establish factors associated with exposure to fringe or radical content. Next, respondents who intentionally accessed fringe or radical content were compared with those who had been exposed to the content unintentionally, to identify factors related to intentional access as distinct from unintentional exposure. Finally, respondents who shared fringe or radical content (irrespective of whether they had seen it intentionally) were compared with respondents who had seen it but did not share it with others online.

## Limitations

Proportional quota-based sampling is a non-probability sampling method. While the sample is large and representative of the spread of the Australian population according to key demographic characteristics, we must be cautious about generalising the results to the wider Australian population.

There are advantages and disadvantages to relying on respondents to assess whether the content they had accessed was fringe, radical or extreme. While we provided guidance to respondents to help them determine whether they had accessed or shared fringe or radical content or participated in a group that promoted fringe or radical views online, respondents may have had different ideas about whether the content was fringe or radical. It is a subjective assessment that is probably linked to their own views around what constitutes mainstream content or the behaviour of most people online. However, relying on respondents to define the content made it possible to measure self-reported exposure and access without collecting information about the content respondents had encountered and without making subjective decisions on whether that content met a pre-determined definition of fringe or radical. Additionally, some respondents may have been reluctant to share information about the content they had accessed or shared, and whether they had done so intentionally, particularly where that content may have been illegal or harmful in nature.

Finally, the distribution of radical content increased during the COVID-19 pandemic (Waldek, Droogan & Ballsun-Stanton 2022). While we do not expect this to have influenced our findings, it is important to note that there were some changes to the online environment shortly prior to this survey.

# Results

## Exposure to and sharing of fringe or radical content online

Forty-one percent of respondents (*n*=5,402) reported having seen fringe or radical content online in the 12 months prior to the survey (Table 1). Of these, 70.2 percent (*n*=3,792)—28.5 percent of all respondents—reported having seen that content unintentionally. A further 23.2 percent (*n*=1,253)—9.4 percent of all respondents—reported intentionally accessing the content. A small group of respondents who said they had seen fringe or radical content did not disclose whether this exposure was intentional or not (6.6%, *n*=345).

Respondents who said they had seen fringe or radical content online were then asked whether they had shared the content. This was limited to respondents who said they had first accessed the material online (*n*=4,101). One in eight respondents (12.7%, *n*=522) who had seen fringe or radical content said they had also shared it online.

| Table 1: Exposure to and sharing of fringe or radical content online (%) | |
|---|---|
| **Exposure (*n*=13,302)** | |
| Had not seen content | 45.9 |
| Had seen content | 40.6 |
| Did not respond | 13.5 |
| **Intentional or unintentional exposure (*n*=5,401)** | |
| Unintentionally viewed content | 70.2 |
| Viewed content intentionally | 23.2 |
| Did not respond | 6.6 |
| **Sharing (*n*=4,101)** | |
| Did not share content | 83.8 |
| Shared content | 12.7 |
| Did not respond | 3.5 |

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Type of content accessed online

Three in 10 respondents who said they had accessed fringe or radical content online said that at least some of this online content described, depicted or advocated violence (29.9%; Table 2). Respondents who said they had unintentionally accessed fringe or radical content online were more likely than respondents who accessed it intentionally to say that the content depicted violence (30.5% vs 28.5%, $F$=24.7, $p$<0.01), although the rate of exposure was similar overall.

Intentional access to fringe or radical content was more common that unintentional access for every form of content. However, podcasts ($F$=225.1, $p$<0.01), presentations or videos ($F$=178.9, $p$<0.01), news or documentaries ($F$=139.8, $p$<0.01), websites ($F$=107.5, $p$<0.01) and online articles ($F$=83.0, $p$<0.01) yielded the largest differences between groups. Respondents who accessed the content intentionally were significantly more likely than unintentional viewers to say that the content was designed to discuss ideas ($F$=225.1, $p$<0.01), as a call to action ($F$=18.4, $p$<0.01), to generate finance ($F$=17.7, $p$<0.01), to provide training ($F$=32.8, $p$<0.01), to facilitate planning ($F$=7.4, $p$<0.01), or simply for entertainment ($F$=122.6, $p$<0.01) or news purposes ($F$=242.4, $p$<0.01).

Respondents who intentionally accessed fringe or radical content were asked why they had done so. The most common reason given was that they wanted to learn more about the cause, had been following the topic or shared some values with proponents (55.6%). This was followed by a general curiosity, for entertainment reasons, or because they had encountered it while browsing (26.2%). One in six (16.5%) said it was because of perceived government and media bias, manipulation and propaganda and a desire to fact check or seek an alternative view, and one in 10 (10.8%) said they had accessed the content because they saw it being promoted on social media.

**Table 2: Type of fringe or radical content accessed online (%)**

| | Unintentionally viewed ($n$=3,792) | Intentionally viewed ($n$=1,253) |
|---|---|---|
| Content described, depicted or advocated violence | 30.5 | 28.5** |
| **Form in which content was accessed** | | |
| Messages, discussions or posts online | 57.9 | 61.1* |
| Documentaries or news articles | 36.9 | 56.5** |
| Websites | 28.1 | 44.3** |
| Presentations or videos | 23.0 | 43.2 ** |
| Online magazines, articles or books | 14.3 | 25.6** |
| Online gaming | 7.8 | 12.0** |
| Podcasts | 7.1 | 23.3** |
| **Aim of content** | | |
| Explanations, information, discussion and ideas | 38.8 | 66.3** |
| Recruitment | 19.0 | 14.4** |
| Call to action | 32.3 | 40.1** |
| Call for violence | 16.0 | 10.7** |
| Financing or donations | 17.1 | 22.4** |
| Training or instruction | 5.7 | 10.5** |
| Planning or executing actions | 8.9 | 11.6** |
| Entertainment or comedy | 11.7 | 25.3** |
| News and current affairs | 27.6 | 52.2** |
| Propaganda | 1.2 | 1.4 |
| Influence and manipulate | 1.2 | 0.2 |

| Table 2: Type of fringe or radical content accessed online (%) (cont.) | | |
|---|---|---|
| | Unintentionally viewed (*n*=3,792) | Intentionally viewed (*n*=1,253) |
| **Most common reasons for accessing content (on most recent occasion)** | | |
| To learn more about the cause, had been following topic or shared some values with proponents | – | 55.6 |
| Curious, for entertainment or because encountered while browsing | – | 26.2 |
| Government or media are biased and engage in manipulation or propaganda, wanted to fact check | – | 16.5 |
| Content was shared or discussed on social media | – | 10.8 |
| Recommended or discussed by someone they knew | – | 7.0 |
| Disagree with content (no further explanation) | – | 5.3 |
| Participate in discussion or exchange of ideas (no further explanation) | – | 2.8 |
| Other reason | – | 6.3 |
| Don't know or prefer not to say | – | 3.5 |

**statistically significant at $p<0.01$, *statistically significant at $p<0.05$

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Characteristics of individuals who accessed or shared content online

As shown in Table 3, respondents who unintentionally viewed fringe or radical content were more likely than respondents who had not seen such content to be younger (30.8% below the age of 35 vs 26.0%; $F=7.3$, $p<0.01$), male (51.9% vs 45.5%; $F=23.3$, $p<0.01$) and university educated (41.4% vs 38.3%; $F=4.9$, $p<0.01$), and to spend more time online for personal (3.8 vs 3.4 hours; $F=3.4$, $p<0.01$) or work (3.2 vs 2.9 hours; $F=2.7$, $p<0.01$) purposes. Respondents who intentionally viewed fringe or radical content were more likely than respondents who were unintentionally exposed to be university educated (48.7% vs 41.4%; $F=13.7$, $p<0.01$) and male (61.7% vs 51.9%; $F=13.7$, $p<0.01$); however, they were also older (51.7% 50 years or over vs 44.4%; $F=5.6$, $p<0.01$), and less likely to be employed (59.8% vs 63.7%; $F=3.67$, $p<0.05$).

Compared to respondents who did not share fringe or radical content online, respondents who shared this content were younger (36.5% below the age of 35 vs 29.4%; $F=3.0$, $p<0.01$), were less likely to be in a relationship (57.1% vs 64.3%; $F=4.47$, $p<0.01$), were less likely to be employed (61.6% vs 65.6%; $F=2.57$, $p<0.05$), and spent significantly more time online for personal activities (4.6 vs 3.9 hours; $F=1.47$, $p<0.05$).

**Table 3: Characteristics of individuals who accessed, intentionally engaged with and shared fringe or radical content online**

| | Had not seen content (*n*=6,104) | Unintentionally viewed content (*n*=3,792) | Intentionally viewed content (*n*=1,253) | Shared content (*n*=4,101) | |
|---|---|---|---|---|---|
| | | | | Yes (*n*=522) | No (*n*=3,435) |
| **Gender (%)** | | | | | |
| Male | 45.5 | 51.9 | 61.7 | 56.5 | 55.7 |
| Female | 54.4 | 47.5** | 37.4** | 42.1 | 43.7 |
| **Age group (%)** | | | | | |
| 18–24 | 8.9 | 11.9 | 9.4 | 14.1 | 11.5 |
| 25–34 | 17.1 | 18.9 | 14.9 | 22.4 | 17.9 |
| 35–49 | 26.4 | 24.8 | 23.9 | 22.8 | 24.7 |
| 50–64 | 24.3 | 22.2 | 25.6 | 21.9 | 24.2 |
| 65+ | 23.3 | 22.2** | 26.1** | 18.8** | 21.8 |
| **Highest level of education (%)** | | | | | |
| High school | 32.8 | 28.7 | 24.8 | 26.5 | 27.2 |
| Vocational | 28.4 | 29.2 | 25.3 | 29.4 | 27.8 |
| University | 38.3 | 41.4** | 48.7** | 43.2 | 44.4 |
| **Family situation (%)** | | | | | |
| Currently in a relationship | 64.1 | 63.9 | 61.2 | 57.1** | 64.3 |
| Children living at home | 37.7 | 35.7** | 35.2 | 36.9 | 35.8 |
| **Employment (%)** | | | | | |
| Employed | 63.1 | 63.7 | 59.8 | 61.6 | 65.6 |
| Unemployed | 4.2 | 4.1 | 5.5** | 7.4** | 4.1 |
| Other | 32.8 | 32.2 | 34.7 | 30.9 | 30.4 |
| **Internet use (hours per day)** | | | | | |
| Personal | 3.4 | 3.8** | 3.9 | 4.6** | 3.9 |
| Work | 2.9 | 3.2** | 2.9 | 3.1** | 3.6 |

**statistically significant at $p<0.01$

Note: 'Other' employment status refers to those who reported being neither employed nor unemployed, such as retired people, full-time carers or those not working due to health conditions, illness or injury

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Initial exposure to content

Respondents who accessed fringe or radical content intentionally were more likely to have experienced their first exposure online (83.3% vs 80.5%; $F$=36.7, $p$<0.01) and to have sought out and found the content themselves (59.6% vs 40.0%; $F$=23.3, $p$<0.01; Table 4). They were also more likely to have first accessed the content while under the age of 18 (5.2% vs 3.4%; $F$=1.41, $p$<0.01). Those who accessed fringe or radical content intentionally reported doing significantly more frequently than those who were exposed unintentionally ($F$=51.31, $p$<0.01), with nearly five times the rate of weekly exposure (34.9% vs 7.2%).

Respondents were also asked whether they knew if their immediate family or friends had accessed fringe or radical content online in the last 12 months. While a large proportion of respondents did not know, respondents who said they intentionally accessed fringe or radical content online were much more likely to also say they had immediate family members (23.0% vs 9.0%; $F$=32.9, $p$<0.01) and friends (29.3% vs 13.0%; $F$=31.5, $p$<0.01) who had accessed similar content online.

**Table 4: Initial exposure and frequency of access to fringe or radical content online**

| | Unintentionally viewed content ($n$=3,792) | Intentionally viewed content ($n$=1,253) |
|---|---|---|
| **Online or offline initial exposure to fringe or radical content (%)** | | |
| Online | 80.5 | 83.3** |
| Offline | 6.5 | 9.7 |
| **How initial exposure occurred (%)** | | |
| Found by themselves | 40.0 | 59.6** |
| Shared by someone they knew or had met in person | 25.3 | 25.5 |
| Shared by someone they met online | 6.9 | 4.7 |
| **Timing of first exposure (%)** | | |
| Last 30 days | 11.5 | 20.1** |
| Last 12 months (but not last 30 days) | 38.6 | 29.5 |
| More than 12 months ago | 31.0 | 43.6 |
| Did not access online | 5.9 | 0.8 |
| **Age of first exposure to online content** | | |
| Mean (*SD*) | 44.2 (18.8) | 44.9 (19.0) |
| Under 18 years (%) | 3.4 | 5.2** |
| **Recent frequency of exposure (%)** | | |
| At least weekly | 7.2 | 34.9 |
| At least monthly | 22.3 | 32.5 |
| Have not accessed in the past month | 59.3 | 26.1** |
| **Family and friends known to have accessed fringe or radical content online in the last 12 months (%)** | | |
| Immediate family | 9.0 | 23.0** |
| Friends | 13.0 | 29.3** |

**statistically significant at $p$<0.01

Note: Denominators include respondents who did not answer the question; therefore percentages may not total 100

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Platforms used by individuals who accessed content online

Respondents who were unintentionally exposed to fringe or radical content were more likely than respondents who were not exposed to engage with almost all platforms weekly, consistent with the higher amount of time spent online (Table 5). However, compared with those who were unintentionally exposed, respondents who intentionally sought out fringe or radical content were less likely to access Facebook (57.5% vs 68.9%; $F$=52.03, $p$<0.01), Messenger (43.4% vs 48.9%; $F$=11.31, $p$<0.01) and TikTok (12.6% vs 17.8%; $F$=15.84, $p$<0.01) weekly, and more likely to engage with Telegram (5.5% vs 2.5%; $F$=24.01, $p$<0.01), Rumble (2.8% vs 0.5%; $F$=39.75, $p$<0.01), 4chan (1.1% vs 0.5%; $F$=5.12, $p$<0.05), Parler (0.6% vs 0.2%; $F$=6.71, $p$<0.01) and Gab (0.6% vs 0.1%; $F$=11.52, $p$<0.01).

Respondents who shared fringe or radical content online were less likely to engage with Facebook (64.8% vs 68.4%; $F$=3.65, $p$<0.05) weekly than respondents who did not, but more likely to engage with Messenger (51.5% vs 49.6%; $F$=2.61, $p$<0.05) and TikTok (21.6% vs 16.2%; $F$=4.68, $p$<0.01). The most notable differences in weekly platform use related to Telegram (9.4% vs 2.8%; $F$=17.79, $p$<0.01) and Rumble (4.2% vs 0.8%; $F$=13.65, $p$<0.01), with those who shared fringe or radical content much more likely to use these at least weekly.

**Table 5: At least weekly use of specific platforms or apps, by exposure and access to fringe or radical content (%)**

| | Had not seen content ($n$=6,104) | Unintentionally viewed content ($n$=3,792) | Intentionally viewed content ($n$=1,253) | Shared content ($n$=4,101) | |
|---|---|---|---|---|---|
| | | | | Yes ($n$=522) | No ($n$=3,435) |
| Facebook | 62.1 | 68.9** | 57.5** | 64.8** | 68.4 |
| Messenger | 40.6 | 48.9** | 43.4** | 51.5** | 49.6 |
| WhatsApp | 25.6 | 27.5** | 26.5 | 24.3 | 28.5 |
| TikTok | 13.9 | 17.8** | 12.6** | 21.6** | 16.2 |
| Twitter | 9.1 | 15.8** | 17.6 | 16.3 | 17.9 |
| Reddit | 5.8 | 11.6** | 11.7 | 10.8 | 13.4 |
| Discord | 3.9 | 7.5** | 7.0 | 9.6 | 7.9 |
| Signal | 2.1 | 2.9** | 3.8 | 4.6 | 3.1 |
| Telegram | 1.5 | 2.5** | 5.5** | 9.4** | 2.8 |
| WeChat | 0.9 | 1.4** | 1.4 | 2.0 | 1.2 |
| Rumble | 0.2 | 0.5** | 2.8** | 4.2** | 0.8 |
| 4chan | 0.2 | 0.5** | 1.1** | 1.4 | 0.6 |
| Parler | 0.1 | 0.2 | 0.6** | 1.0 | 0.2 |
| Gab | 0.1 | 0.1 | 0.6** | 1.2 | 0.1 |
| 8kun | 0.1 | 0.2 | 0.4 | 0.9 | 0.1 |
| Kiwi Farms | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |

**statistically significant at $p$<0.01

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Platforms used to share fringe or radical content and join groups online

Respondents who reported sharing fringe or radical content (*n*=522) were asked which online platforms they had used to share the content. Most respondents who shared this content did so across social media or communications platforms such as Facebook (30.4%), Messenger (17.4%) or email (11.2%; Table 6). Relatively few respondents identified multiple platforms, suggesting this was the platform respondents most strongly associated with sharing fringe or radical content.

| Table 6: Mainstream platforms used to share fringe or radical content (%) | |
|---|---|
| Facebook | 30.3 |
| Messenger | 17.4 |
| Email | 11.2 |
| Instagram | 5.5 |
| WhatsApp | 4.3 |
| Twitter | 4.2 |
| TikTok | 3.9 |
| Telegram | 2.9 |
| Reddit | 1.4 |
| Snapchat | 1.3 |
| Signal | 0.9 |
| In-person | 2.5 |
| Unspecified messaging app | 14.2 |
| Unspecified social media | 3.9 |
| Unspecified forum | 1.6 |
| Other platform | 5.1 |
| Don't know | 1.9 |

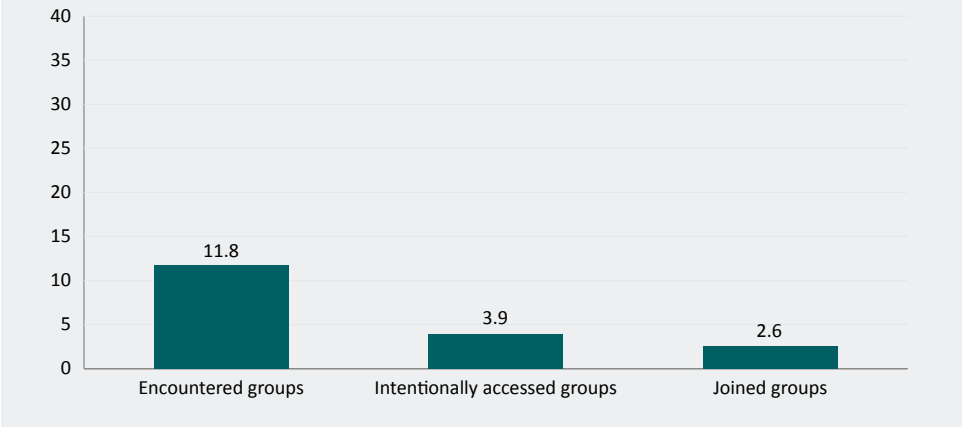
Note: Percentages do not sum to 100 because some respondents indicated sharing content across more than one platform

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Access to fringe or radical online groups

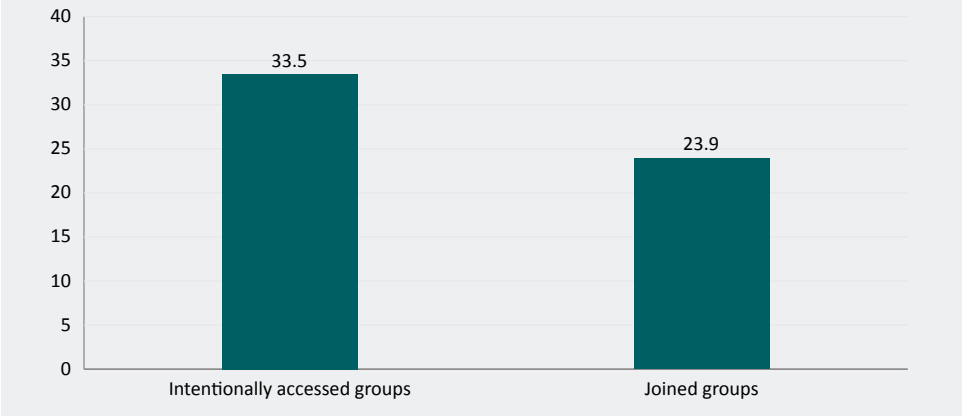
Around one in eight respondents (11.8%, *n*=1,563) had encountered a group online that promoted fringe or radical views, 3.9 percent (*n*=524) intentionally accessed these groups to view the content and 2.6 percent (*n*=346) reported that they had joined one of these groups (Figure 1). One in four respondents who said they had encountered a group online that promoted fringe or radical views said they went on to join a group (23.9%, *n*=374; Figure 2).

**Figure 1: Proportion of total sample who encountered and participated in online groups promoting fringe or radical views (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

**Figure 2: Proportion of respondents who encountered online groups promoting fringe or radical views who went on to access or join those groups (%)**




Source: Survey of social and political attitudes in Australia, 2022 [computer file]

Among respondents who had viewed fringe or radical content unintentionally in the last 12 months, 19.4 percent (*n*=735) had encountered a fringe or radical group and 3.5 percent had intentionally accessed the group. Only 2.6 percent of respondents who had been exposed to fringe or radical content online reported intentionally accessing and joining a fringe or radical group and 2.2 percent had invited others to join the group. It is possible that these individuals joined the group (such as a group on a social media platform) before they became aware of the nature of the content being shared by that group.

A significantly higher proportion of respondents who had intentionally accessed fringe or radical content had encountered (34.1% vs 19.4%; $F$=39.9, $p$<0.01), intentionally accessed (24.9% vs 9.1%; $F$=69.0, $p$<0.01), joined (16.4% vs 7.9%; $F$=18.6, $p$<0.01) and invited others to join these groups (2.8% vs 1.5%; $F$=56.2, $p$<0.01), when compared with respondents who had viewed the content unintentionally. Similarly, a higher proportion of respondents who shared fringe or radical content encountered (45.1% vs 22.3%; $F$=32.3, $p$<0.01), intentionally accessed (68.1% vs 25.2%; $F$=28.4, $p$<0.01), joined (56.7% vs 14.1%; $F$=27.5, $p$<0.01) and invited others to join these groups (23.7% vs 2.1%; $F$=58.3, $p$<0.01), when compared with respondents who had not shared such content online. Respondents who shared fringe or radical content were the most likely to have joined (56.7%, $n$=296) and invited others to join groups promoting fringe or radical views (23.7%, $n$=124). Put another way, 42.9 percent of respondents who had intentionally accessed a fringe or radical group, and 52.9 percent of respondents who had joined a fringe or radical group, said they had shared fringe or radical content with others in the 12 months prior to the survey. Among those who had intentionally accessed fringe or radical content, respondents who had joined a fringe or radical group were also much more likely to share content than respondents who were not a member (68.1% vs 27.7%; $F$=11.3, $p$<0.01).

The platform most commonly used to access a group promoting fringe or radical views—regardless of the manner in which respondents viewed or shared fringe or radical content—was Facebook.

**Table 7: Participation in online groups promoting fringe or radical views (%)**

| | Unintentionally viewed content ($n$=3,792)[a] | Intentionally viewed content ($n$=1,253)[a] | Shared content ($n$=4,101) | |
|---|---|---|---|---|
| | | | Yes ($n$=522)[b] | No ($n$=3,435)[b] |
| Encountered groups | 19.4 | 34.1** | 45.1** | 22.3 |
| Intentionally accessed groups | 9.1 | 24.9** | 68.1** | 25.2 |
| Joined groups | 7.9 | 16.4** | 56.7** | 14.1 |
| Invited others to join the group | 1.5 | 2.8** | 23.7** | 2.1 |
| **Platform used to access a group** | | | | |
| Facebook | 49.9 | 36.0* | 39.3 | 44.4 |
| Social media (not specified) | 3.7 | 9.9 | 12.1 | 5.5* |
| Reddit | 13.9 | 10.4 | 5.11 | 17.9** |
| TikTok | 5.3 | 2.4 | 4.3 | 2.1 |
| Telegram | 3.9 | 5.4* | 4.9 | 5.0 |
| Email | 1.9 | 3.9 | 3.9 | 2.5 |
| Forums | 2.8 | 8.8 | 5.3 | 8.9 |
| Discord | 0 | 1.4 | 0.0 | 0.1 |
| 4chan | 2.9 | 3.9 | 0.1 | 6.8* |
| YouTube | 5.1 | 6.8 | 6.9 | 7.8 |
| Instagram | 2.2 | 4.0 | 4.1 | 3.8 |

**statistically significant at $p$<0.01, *statistically significant at $p$<0.05

a: Excludes 356 respondents who did not answer questions about whether they had accessed fringe or radical content intentionally or unintentionally

b: Excludes 144 respondents who did not answer questions about whether they had shared fringe or radical content

Source: Survey of social and political attitudes in Australia, 2022 [computer file]

## Discussion

Respondents who were unintentionally exposed to fringe or radical content, or who intentionally accessed or shared such content online, tended to be university educated men who were younger than other respondents and spent a greater amount of time online. This profile is consistent with other research using community samples (Costello et al. 2016). The respondents who accessed radical content intentionally were more likely to spend time on all types of platforms, with the largest differences observed between podcasts, presentations or videos, news or documentaries, websites and online articles. They were also more likely to be under the age of 18 when they first saw radical content, to seek out the content themselves, to access the content at least weekly and to have been accessing the content for more than 12 months, potentially indicating a greater commitment to the online community in which they participate (Scrivens 2021). Pivotally, those who intentionally accessed the content used mainstream social media platforms and messaging apps such as Facebook, TikTok, Messenger and WhatsApp less frequently than other respondents, but used platforms such as Telegram, Rumble and 4chan more frequently. They were also more likely to say their immediate family members and friends had also accessed fringe or radical content online in the 12 months prior to the survey, highlighting the intersection between offline and online networks (Gill et al. 2017).

When respondents were asked why they had accessed fringe or radical content, the most common reason was to learn more about a cause. However, a sizeable minority specifically sought out the content because of perceived government or media bias, manipulation or propaganda. This suggests that many people who sought out the content did so because it aligned with their world views, at least to some degree. One in 10 said the main reason for accessing the content was that they had seen it promoted on social media, reinforcing the prominent role that platforms play in shaping people's online behaviour, including in ways that may be harmful (Droogan, Waldek & Blackhall 2018).

Conversely, respondents who were exposed to fringe or radical content unintentionally were much less likely to have found it themselves and, while they accessed the content less frequently, nearly 90 percent experienced their first exposure online. They spent significantly more time online for both work and personal reasons and, during their time online, were more likely than people who were not exposed to fringe or radical content to be frequent users of unmoderated platforms such as Signal, Telegram, WeChat, Rumble and 4chan. They were also the group most likely to report exposure to content that depicted or advocated violence. It is possible that individuals who intentionally accessed fringe or radical content, by virtue of seeking it out, were better able to control the type of content to which they were exposed, whereas those who were unintentionally exposed could not, leading to a greater chance of exposure to violent content. The latter group were also more sensitive to the aims of the content and were more likely to identify it as calling for violence or as attempting to recruit individuals to a certain ideology.

Respondents who shared fringe or radical content online were younger than those who did not, more likely to be unemployed and spent more personal time online. Similar to those who accessed radical content intentionally, respondents who shared it spent less time on mainstream social media platforms than those who did not share, with the exception of Messenger and TikTok. Notably, those who shared fringe or radical content were significantly more likely than those who did not share to use Telegram, WeChat and Rumble weekly. Although they may have spent less time on mainstream social media, this group reported sharing fringe or radical content across conventional platforms such as Facebook, Messenger, email, Instagram, WhatsApp and Twitter. It is possible that respondents who shared radical content accessed the content from unmoderated platforms and shared it across moderated platforms to individuals who may not have been seeking it out, potentially brokering unintentional exposure on mainstream platforms. The transition of content from unmoderated to mainstream platforms supports the increased normalisation of the content. As this content is increasingly shared across mainstream platforms, opportunities for exposure expand to a wider group, normalising and further supporting the sharing of the content and ultimately increasing the risk of exposure among those who may be vulnerable to radicalisation.

Nearly one in eight respondents had encountered online groups that promoted fringe or radical views. One-third of these respondents accessed and engaged with the groups they had encountered and nearly one-quarter reported joining a group. Individuals who intentionally accessed or shared fringe or radical content accessed and joined online groups at a much higher rate than respondents who encountered the content unintentionally. Around half of those respondents who had shared content online were also member of a group promoting fringe or radical content. Further, two-thirds of respondents who had intentionally accessed fringe or radical content and joined a group online shared the content with other internet users.

While this study has established a strong relationship between intentionally accessing and sharing content and participation in online groups, we cannot be certain about the direction of this relationship. However, it is clear that a large proportion of survey respondents had been exposed to fringe or radical content online on mainstream platforms, many of whom had sought it out intentionally. Those who intentionally sought out this content were more likely than those who were exposed unintentionally to participate in an online group promoting fringe or radical content. Being a member of a group promoting fringe or radical content was clearly associated with increased sharing of that content with other internet users. These individuals act as brokers of fringe and radical content online and of networks that facilitate exposure to this content. Prior research has suggested that, rather than being viewed as simple exposure, online interactions with radical content could at times be considered sophisticated and purposeful (Waldek, Droogan & Lumby 2021). Our findings suggest that active exposure to and sharing of fringe or radical content can be a sophisticated, non-trivial exercise, laying the foundation for the spread of this content online.

Results highlight the importance of removing fringe or radical content from online platforms to reduce exposure to this content. This is especially true of mainstream social media platforms, which were the primary platforms on which radical content was shared and the platforms most commonly used by those unintentionally exposed to the content. This approach may result in reduced downstream opportunities for the sharing of fringe or radical content and, as a result, fewer people joining and recruiting others to join online groups that promote this content.

## References

*URLs correct as at August 2024*

Australian Bureau of Statistics (ABS) 2023. National, state and territory population, September 2022. https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/sep-2022

Bastug MF, Douai A & Akca D 2020. Exploring the "demand side" of online radicalization: Evidence from the Canadian context. *Studies in Conflict & Terrorism* 43(7): 616–637. https://doi.org/10.1080/1057610X.2018.1494409

Broadhurst R, Woodford-Smith H, Maxim D, Sabol B, Orlando S, Chapman-Schmidt B & Alazab M 2017. *Cyber terrorism: Research review*. Canberra: Australian National University Cybercrime Observatory. https://doi.org/10.2139/ssrn.2984101

Costello M, Hawdon J, Ratliff T & Grantham T 2016. Who views online extremism? Individual attributes leading to exposure. *Computers in Human Behavior* 63: 311–320. https://doi.org/10.1016/j.chb.2016.05.033

Droogan J & Waldek L 2018. Should we be afraid of cyber-terrorism? *International Journal of Electronic Security and Digital Forensics* 10(3): 242–254. https://doi.org/10.1504/IJESDF.2018.093017

Droogan J, Waldek L, Ballsun-Stanton B & Hutchinson J 2022. *Mapping a social media ecosystem: Outlinking on Gab & Twitter amongst the Australian far-right milieu*. Washington, DC: Resolve Network. https://doi.org/10.37805/remve2022.6

Droogan J, Waldek L & Blackhall R 2018. Innovation and terror: An analysis of the use of social media by terror-related groups in the Asia Pacific. *Journal of Policing, Intelligence and Counter Terrorism* 13(2): 170–184. https://doi.org/10.1080/18335330.2018.1476773

Frischlich L 2021. #Dark inspiration: Eudaimonic entertainment in extremist Instagram posts. *New Media & Society* 23(3): 554–577. https://doi.org/10.1177/1461444819899625

Frissen T 2021. Internet, the great radicalizer? Exploring relationships between seeking for online extremist materials and cognitive radicalization in young adults. *Computers in Human Behavior* 114(1): 106549. https://doi.org/10.1016/j.chb.2020.106549

Gaudette T, Scrivens R & Venkatesh V 2022. The role of the internet in facilitating violent extremism: Insights from former right-wing extremists. *Terrorism and Political Violence* 34(7): 1339–1356. https://doi.org/10.1080/09546553.2020.1784147

Gill P, Corner E, Conway M, Thornton A, Bloom M & Horgan J 2017. Terrorist use of the internet by the numbers: Quantifying behaviors, patterns, and processes. *Criminology & Public Policy* 16(1): 99–117. https://doi.org/10.1111/1745-9133.12249

Hassan G et al. 2018. Exposure to extremist online content could lead to violent radicalization: A systematic review of empirical evidence. *International Journal of Developmental Science* 12(1–2): 71–88. https://doi.org/10.3233/DEV-170233

Nienierza A, Reinemann C, Fawzi N, Riesmeyer C & Neumann K 2021. Too dark to see? Explaining adolescents' contact with online extremism and their ability to recognize it. *Information, Communication & Society* 24(9): 1229–1246. https://doi.org/10.1080/1369118X.2019.1697339

Reeve Z 2023. Repeated and extensive exposure to online terrorist content: Counter-Terrorism Internet Referral Unit perceived stresses and strategies. *Studies in Conflict & Terrorism* 46(6): 888–912. https://doi.org/10.1080/1057610X.2020.1792726

Scrivens R 2021. Exploring radical right-wing posting behaviors online. *Deviant Behavior* 42(11): 1470–1484. https://doi.org/10.1080/01639625.2020.1756391

Waldek L, Droogan J & Ballsun-Stanton B 2022. *Online far right extremist and conspiratorial narratives during the COVID-19 pandemic.* Sydney: Macquarie University. https://doi.org/10.5281/zenodo.5732611

Waldek L, Droogan J & Lumby C 2021. *Feeling terrified? The emotions of online violent extremism*. Cambridge University Press: Cambridge. https://doi.org/10.1017/9781108886369

Walther S & McCoy A 2021. US extremism on Telegram: Fueling disinformation, conspiracy theories, and accelerationism. *Perspectives on Terrorism* 15(2): 100–124. https://www.jstor.org/stable/27007298

Whittaker J 2021. The online behaviors of Islamic state terrorists in the United States. *Criminology & Public Policy* 20(1): 177–203. https://doi.org/10.1111/1745-9133.12537

Wolbers H, Dowling C, Cubitt T & Kuhn C 2023. Understanding and preventing internet-facilitated radicalisation. *Trends & issues in crime and criminal justice* no. 673. Canberra: Australian institute of Criminology. https://doi.org/10.52922/ti77024

Wolfowicz M, Litmanovitz Y, Weisburd D & Hasisi B 2021. Cognitive and behavioral radicalization: A systematic review of the putative risk and protective factors. *Campbell Systematic Reviews* 17(3): e1174. https://doi.org/10.1002/cl2.1174

This work was originally published as *Trends & issues in crime and criminal justice* no. 705. It is available online at **https://doi.org/10.52922/ti77734**.

# 6. The overlap between viewing child sexual abuse material and fringe or radical content online

Timothy Cubitt, Anthony Morgan and Rick Brown

## Introduction

In recent years, online dissemination of radical material by extremist groups has increased, resulting in exposure to a wider audience (Broadhurst et al. 2017; Droogan & Waldek 2018). Simultaneously, the growth of the internet and related technologies has increased the availability of child sexual abuse material (CSAM; Holt et al. 2020; WePROTECT Global Alliance 2021; Westlake 2020). Access to both types of content has proliferated in recent years; however, they have typically been viewed as separate problems involving different people. This study examines the characteristics and behaviours of Australians who access either CSAM or radical content online and compares them with individuals who access both.

### The problem of radical content online

Radical content accessed through online domains is an important risk factor for radicalisation (Australian Security Intelligence Organisation 2023; Wolfowicz, Hasisi & Weisburd 2022) but not necessarily a standalone cause (Gill et al. 2017), resulting in a recent focus on how exposure occurs and who is at the greatest risk (Cubitt & Morgan 2024; Wolbers et al. 2023). Platforms with less moderation, such as 4chan, 8chan and Telegram, allow greater reach and accessibility of radical content (Hassan et al. 2018; Nienierza et al. 2021), with encrypted communication also supporting peer-to-peer sharing (Nuraniyah 2019).

Recent research has shown that a large proportion of computer users are exposed to online content they describe as fringe, radical or extreme, a significant minority of whom intentionally accessed the content (Cubitt & Morgan 2024). Exposure to this type of content, especially when it is intentionally sought, is associated with extremist attitudes and behaviours (Hassan et al. 2018; Wolfowicz, Hasisi & Weisburd 2022). The internet is an important component of modern radicalisation pathways, with up to 60 percent of extremists found to have radicalised primarily online or to have had considerable online engagement with radical content in combination with offline influences (Hamid & Ariza 2022; Kenyon, Binder & Baker-Beall 2023).

Content detection and removal, account suspension, identity validation and counter-narrative approaches have been supported by government agencies, and implemented by the private sector, to reduce the proliferation of radical content online (Wolbers et al. 2023). However, the problem persists, on mainstream social media platforms as well as those featuring less moderation (Cubitt & Morgan 2024).

## Online child sexual abuse material

CSAM is also widely available online (Broadhurst 2019; Westlake 2020). While it is a significant problem on the darknet (Johnson & Patel 2019), CSAM is also found in indexed content on surface web search engines (Westlake 2020); searches for adult pornography (Morgan & Lambie 2019); peer-to-peer networks (Wolak, Liberatore & Levine 2014); and image boards, file-storage sites and widely used social media sites (Brown 2022). While internet relay chat and email were once common methods for privately sharing CSAM, encrypted messaging apps have become popular for sharing content (Balfe et al. 2015; Bursztein et al. 2019; Brown 2022). Among the pathways into CSAM offending, the first exposure is an important point in the onset of offending (Hunn et al. 2023; Wortley & Smallbone 2012), with impulsivity and risk-taking behaviour an established contributory factor (Seto & Ahmed 2014).

International estimates of the prevalence of viewing CSAM have ranged from 2.2 percent to 4.6 percent of adults having ever seen such content (Brown 2023). In Australia, Salter et al. (2023) estimated that 2.5 percent of men had ever viewed CSAM. Using a narrower definition, Brown (2023) estimated that 0.8 percent of all adults (and 1.1% of men) had intentionally viewed CSAM in the previous year.

There is growing acknowledgement of a nexus between CSAM and radical online content. For example, in January 2023 a member of an online neo-Nazi group was sentenced in a United Kingdom court for possessing CSAM (see https://www.bbc.com/news/uk-wales-64404704). In September 2023, the US Federal Bureau of Investigation released a public service announcement warning of violent online groups involved in child sexual extortion, with CSAM being produced and shared with other group members (Federal Bureau of Investigation 2023). Further, in early 2024, the Global Network on Extremism and Technology noted that there may be an intersection between terrorism, violent extremism and child sexual exploitation (Argentino, Barrett & Tyler 2024).

Despite the fact that a large proportion of both CSAM and radical content exists on the surface web and social media, to date there has been no research into individuals who engage with both radical content and CSAM. This research was prompted by observations from law enforcement and terrorism professionals that there may be individuals prone to accessing both types of content online (Argentino, Barrett & Tyler 2024; Federal Bureau of Investigation 2023), as well as separate consultation with law enforcement and academia undertaken by the authors. Using a large-scale survey of Australian internet users, this study examines the characteristics and behaviours of respondents who viewed CSAM and fringe or radical content online, with a focus on respondents who viewed both types of content.

# Method

## Sample

This research uses data from a large sample of online Australians surveyed about their political and social beliefs (*n*=13,302). Conducted by Roy Morgan Research in late 2022, the survey was sent to panels of individuals aged at least 18 years who had voluntarily joined to receive incentives in exchange for completing surveys. Participants were first recruited from Roy Morgan's Single Source survey panel, which uses a cluster sampled, face-to-face survey approach. Proportional quota sampling ensured that the sample was reflective of the spread of people living in Australia, while the data were weighted by age, sex and usual place of residence. Additional random iterative method weights were applied to ensure representativeness of the sample by education level, internet and social media use. Quotas were based on the Australian adult population stratified by age, sex and usual place of residence, based on Australian Bureau of Statistics (2023) population data. The raw completion rate for invitations sent to this panel was 7.1 percent. The survey took respondents an average of 23.8 minutes to complete. Poor quality responses and duplicates were removed from the final sample.

Respondents were asked whether, in the last 12 months, they had seen political, ideological or religious content online that mainstream society or the government might describe as fringe, unorthodox or radical (referred to as 'fringe or radical content'). This included online messages, posts, or discussions; videos they had watched, shared on bulletin boards, or attached to text or email messages (excluding fictional films); and books, magazines, articles, or other written material, downloaded to print or received attached to a message such as an email. These may have included political or religious content such as speeches, sermons, manifestos, policies, ideas or information designed to influence respondents' views on world events towards an identifiable ideology or issue. The content may have been viewed intentionally or unintentionally. They were also asked whether, in the last 12 months, they had encountered online sexually explicit material of people who are or look under the age of 18. Using the same dataset, Brown (2023) focused specifically on those who had not only viewed CSAM but done so intentionally. The present study includes all who had viewed CSAM in the previous 12 months, regardless of whether it was intentional. The choice was made to include those who accessed each type of content both intentionally and unintentionally for analysis. This was for several reasons: to ensure that the groups analysed were similar; to acknowledge that viewing these types of content can occur for many reasons, some of which may be situational; and to allow inferences to be made in relation to the totality of the sample, rather than only a smaller subset.

Based on their responses to these questions, respondents were categorised into one of four groups:

- respondents who said they had not viewed CSAM or fringe or radical content in the 12 months prior to the survey;
- respondents who said they had viewed CSAM but not fringe or radical content (CSAM only group);
- respondents who said they had viewed fringe or radical content but not CSAM (fringe or radical content group); and
- respondents who said they had viewed both CSAM and fringe or radical content (overlap group).

Respondents who did not answer one or both questions about viewing CSAM or fringe or radical content were excluded from the comparative analysis; however, they are included in the overall prevalence estimates.

## Analytical approach

The analysis proceeded in two stages. The first stage involved bivariate analyses exploring the differences between the four groups in their demographic characteristics, regular use of the internet and online platforms, reported criminal justice contacts as a juvenile and as an adult, and diagnosed mental illness. We note that, while we asked about regular use of platforms, we did not ask whether respondents accessed any type of content on these specific platforms. Rather, these questions related to their regular online activities. Each of the three groups who viewed one or both types of content were initially compared with respondents who had viewed neither type of content. Next, the CSAM group was separately compared with the fringe or radical content group and the overlap group. Finally, the fringe or radical content group was compared with the overlap group.

Chi-square tests of association were used to measure the statistical significance of any differences observed between groups. Because this study uses weighted survey data, we report an *F* statistic rather than the $\chi^2$ value. More specifically, the Pearson's chi-square statistic is turned into an *F* statistic with non-integer degrees of freedom by using a second-order Rao and Scott (1984, 1981) correction. The Rao and Scott correction is used when survey weights are applied to a complex sample with the intention of improving representativeness of the population studied. Where a $\chi^2$ value reports on unweighted data, and in this case may produce a false evaluation of statistical significance, the reported *F* statistic is corrected for the weighted data and therefore more reliable in this circumstance.

## Limitations

While the sample is large and representative of the spread of the Australian population according to key demographic characteristics, caution should be exercised in generalising the results to the wider population. Caution should also be taken in interpreting group membership. While we provided guidance to respondents to help them determine whether they had viewed fringe or radical content online, including specific examples, respondents may have had different ideas about whether the content was fringe or radical. In this study, individuals were categorised into one of four groups based on whether or not they had viewed CSAM and/or fringe or radical content; however, we did not ask the extent to which content was viewed. This meant it was not possible to assess whether prolific consumers of either type of content were different to those who view it less frequently. We also did not examine intentionality in this analysis, which means we do not know whether CSAM or fringe or radical content were viewed accidentally or on purpose. Nor did we ask those who viewed both types of content which was their primary interest and which was secondary. This limited our ability to draw inferences about the intention of the overlap group in viewing both types of content. Further, we did not ask about any other types of content sought. It is possible that these individuals viewed a suite of content types and we were only able to observe two. Finally, some respondents may have been reluctant to share information about the content they had viewed. Given the proportion of the sample who responded 'Don't know' or 'Prefer not to say' when asked whether they had viewed either CSAM or fringe or radical content, it is likely that some respondents preferred not to provide this information.

# Results

## Viewing CSAM or fringe or radical content online

Overall, 39.9 percent of respondents (*n*=5,304) said they had viewed either CSAM or fringe or radical content online in the 12 months prior to the survey. About a third of respondents (35.8%) had viewed fringe or radical content but not CSAM (*n*=4,768), 1.1 percent of respondents (*n*=150) had viewed CSAM but not fringe or radical content, and 2.9 percent of respondents (*n*=386) said they had viewed both types of content. Around one in six respondents (16.4%, *n*=2,176) did not answer at least one of the questions, and for each group these respondents are reported as 'unknown' in Table 1.

The majority (64.7%) of respondents who viewed CSAM had also viewed fringe and radical content online. Conversely, relatively few respondents who had viewed fringe and radical content online had viewed CSAM (7.1%).

**Table 1: Respondents who had viewed CSAM or fringe or radical content in the 12 months prior to the survey (%)**

| | | Viewed fringe or radical content in the past 12 months | | | |
|---|---|---|---|---|---|
| | | Yes | No | Unknown | Total |
| Viewed CSAM in the past 12 months | Yes | 2.9 (*n*=386) | 1.1 (*n*=150) | 0.5 (*n*=60) | 4.5 (*n*=596) |
| | No | 35.8 (*n*=4,768) | 43.8 (*n*=5,822) | 11.8 (*n*=1,569) | 91.4 (*n*=12,159) |
| | Unknown | 1.9 (*n*=247) | 1.0 (*n*=135) | 1.2 (*n*=165) | 4.1 (*n*=547) |
| | Total | 40.6 (*n*=5,402) | 45.9 (*n*=6,105) | 13.5 (*n*=1,795) | 100.0 (*n*=13,302) |

Note: Unknown includes respondents who said they did not know or preferred not to say

Source: Survey of social and political attitudes in Australia 2022 [computer file]

## Sociodemographic characteristics of respondents

Respondents who viewed either one or both types of content were significantly different from respondents who did not across almost all demographic variables. The CSAM ($F$=3.9, $p$<0.05) and overlap groups ($F$=21.1, $p$<0.01) were significantly younger than those who viewed neither. Among those who viewed fringe or radical content, the spread of ages was significantly different to those who did not view either type of content. These respondents were significantly more likely than the group who viewed neither type of content to be under 35, or over 65 years ($F$=5.9, $p$<0.01).

The overlap group was also significantly younger than the CSAM ($F$=3.3, $p$<0.05) and the fringe or radical content ($F$=14.6, $p$<0.01) groups, with a higher proportion of respondents aged 18 to 34 years (40.7% vs 29.7% and 28.5%, respectively). Compared with the group who viewed neither type of content (44.8%), the CSAM (54.8%) and fringe or radical content (53.2%) groups were both significantly more likely to be male ($F$=7.7, $p$<0.01 and $F$=28.1, $p$<0.01 respectively), while the overlap group featured the highest proportion of male respondents (63.5%; $F$=28.2, $p$<0.01). Each of the CSAM group (4.3%; $F$=3.8, $p$=0.05), the fringe or radical content group (2.6%; $F$=5.9, $p$<0.05) and the overlap group (7.5%; $F$=40.6, $p$<0.01) featured a higher proportion of Indigenous respondents than the group which viewed neither type of content (1.9%).

There were relatively few differences between the four groups in terms of employment, education and income. More than one in five respondents (21.4%) in the overlap group had a history of military service, significantly higher than those who had viewed neither type of content (7.4%; $F$=21.4, $p$<0.01) and the fringe or radical content group (9.8%; $F$=22.4, $p$<0.01).

**Table 2: Sociodemographic characteristics of respondents, by group (%)**

| | Viewed neither CSAM nor fringe or radical content (*n*=5,822) | Viewed only CSAM (*n*=150) | Viewed only fringe or radical content (*n*=4,768) | Viewed both CSAM and fringe or radical content (*n*=386) |
|---|---|---|---|---|
| **Age** | | | | |
| 18–34 | 25.6 | 29.7 | 28.5 | 40.7 |
| 35–64 | 50.6 | 56.7 | 47.5 | 44.4 |
| 65+ | 23.8 | 13.6** | 24.1**## | 14.9**##xx |
| **Gender** | | | | |
| Male | 44.8 | 54.8 | 53.2 | 63.7 |
| Female | 55.1 | 44.0** | 46.2** | 35.4**xx |
| Indigenous | 1.9 | 4.3* | 2.6** | 7.5**xx |
| **Annual income ($)** | | | | |
| 0–18,200 | 11.5 | 15.1 | 11.6 | 13.8 |
| 18,201–45,000 | 24.2 | 18.2 | 24.3 | 25.3 |
| 45,001–120,000 | 43.2 | 40.9 | 43.3 | 40.1 |
| 120,001–180,000 | 8.8 | 10.3 | 10.3 | 10.4 |
| 180,001+ | 3.4 | 7.4 | 3.6** | 3.5* |
| **Highest level of education** | | | | |
| High school | 32.7 | 37.1 | 28.0 | 28.7 |
| Vocational | 24.5 | 24.5 | 28.2 | 31.5 |
| University | 38.3 | 38.4 | 43.1** | 39.2 |
| **Current relationship** | 64.3 | 63.2 | 63.1 | 62.8 |
| **Employed** | 62.6 | 73.8** | 62.3 | 64.6 |
| **Military service** | 7.4 | 12.5 | 9.8** | 21.4**xx |

For comparisons with respondents who did not view either CSAM or fringe or radical content, ** $p<0.01$ and * $p<0.05$.

For comparisons with respondents who only viewed CSAM, ## $p<0.01$ and # $p<0.05$.

For comparisons with respondents who only viewed fringe or radical content, xx $p<0.01$ and x $p<0.05$.

Source: Survey of social and political attitudes in Australia 2022 [computer file]

## Internet use

All three groups who viewed one or both types of content spent significantly more time on the internet for personal purposes than the group who had not (CSAM 4.1 vs 3.4 hours: $F$=2.2, $p$<0.01; fringe or radical 3.8 vs 3.4 hours: $F$=3.3, $p$<0.01; overlap 4.6 vs 3.4 hours: $F$=4.6, $p$<0.01). The same was also true of time spent on the internet for work purposes (CSAM 3.4 vs 2.9 hours: $F$=7.2, $p$<0.01; fringe or radical 4.1 vs 3.4 hours: $F$=2.4, $p$<0.01; overlap 4.1 vs 3.4 hours: $F$=5.1, $p$<0.01). The overlap group spent the most time online for both personal and work reasons. There were important differences between groups in whether they used certain platforms frequently, defined as at least weekly use. The group who viewed neither type of content were the least likely to frequently use all platforms other than Signal. There were some differences between the CSAM group and those who viewed neither CSAM nor fringe or radical content, with the CSAM group more likely to frequently use TikTok (24.7% vs 13.6%; $F$=12.8, $p$<0.01), Twitter (14.7% vs 8.8%; $F$=5.2, $p$<0.05) and Reddit (9.9% vs 5.6%; $F$=3.9, $p$<0.05). This group was the most likely of all four groups to frequently use TikTok, with almost one in four using the platform at least weekly.

When compared with respondents who viewed neither type of content, the fringe or radical content group were more likely to frequently use all platforms other than WhatsApp and WeChat. The only difference between the CSAM and fringe or radical content groups was the frequent use of TikTok, with a larger proportion of the CSAM group using TikTok at least weekly (24.7% vs 16.2%; $F$=6.4, $p$<0.05).

There were important differences between the overlap group and other groups in the platforms used. Fifty-six percent of respondents in the overlap group used Facebook at least weekly, a smaller proportion than in every other group (neither content: 62.0%, $F$=4.1, $p$<0.05; CSAM: 66.6%, $F$=4.4, $p$<0.05; fringe or radical: 66.9%, $F$=15.8, $p$<0.01). More than one in five of those in the overlap group (21.9%) used TikTok at least weekly, significantly more than in the groups who viewed neither (13.6%; $F$=17.1, $p$<0.01) or fringe or radical content (16.2%; $F$=6.8, $p$<0.01). A similar proportion of the overlap group used Twitter at least weekly, which was also higher than among those who viewed neither type of content (8.8%; $F$=62.1, $p$<0.01) and fringe or radical content only (15.3%; $F$=9.8, $p$<0.01).

Importantly, the use of encrypted communications differed between the groups. The proportions of overlap group respondents who used Discord and Telegram weekly were roughly double that of respondents from the fringe or radical content group (Discord: 12.8% vs 6.6%, $F$=16.7, $p$<0.01; Telegram: 6.3% vs 1.9%, $F$=103.0, $p$<0.01). The overlap group used Signal at four times the rate of the CSAM group (4.4% vs 1.1%; $F$=4.24, $p$<0.05) and nearly twice the rate of the fringe and radical group (4.4% vs 2.9%), although this difference was not statistically significant. Frequent use of Discord was more than three times as common among the overlap group compared with respondents who did not view either type of content (12.8% vs 3.7%; $F$=58.7, $p$<0.01). While the remaining platforms were accessed less frequently, the proportion of respondents active on them was higher among the overlap group than among those who engaged with neither type of content or fringe or radical content only.

**Table 3: Internet use by respondents, by group**

| | Viewed neither CSAM nor fringe or radical content (*n*=5,822) | Viewed only CSAM (*n*=150) | Viewed only fringe or radical content (*n*=4,768) | Viewed both CSAM and fringe or radical content (*n*=387) |
|---|---|---|---|---|
| **Internet hours for personal use** | 3.4 | 4.1** | 3.8**## | 4.6**xx |
| **Internet hours for work** | 2.9 | 3.4** | 3.1**## | 3.8**xx |
| **Platforms used at least once per week (%)** | | | | |
| Facebook | 62.0 | 66.6 | 66.9** | 56.3*#xx |
| Messenger | 40.6 | 42.7 | 47.7** | 45.2 |
| WhatsApp | 25.6 | 24.6 | 27.3 | 25.8 |
| TikTok | 13.6 | 24.7** | 16.2**## | 21.9**xx |
| Twitter | 8.8 | 14.7* | 15.3** | 21.8**xx |
| Reddit | 5.6 | 9.9* | 10.7** | 14.4**xx |
| Discord | 3.7 | 6.1 | 6.6** | 12.8**xx |
| Telegram | 1.5 | 1.9 | 3.1** | 6.3**xx |
| Signal | 2.1 | 1.1 | 2.9** | 4.4**# |
| WeChat | 0.9 | 1.2 | 1.2 | 2.4**xx |
| Rumble | 0.2 | 0.8 | 1.0** | 2.4**xx |
| 4chan | 0.1 | 0.7 | 0.6** | 2.6**xx |

For comparisons with respondents who did not view either CSAM or fringe or radical content, ** $p<0.01$ and * $p<0.05$

For comparisons with respondents who only viewed CSAM, ## $p<0.01$ and # $p<0.05$

For comparisons with respondents who only viewed fringe or radical content, xx $p<0.01$ and x $p<0.05$

Source: Survey of social and political attitudes in Australia 2022 [computer file]

## Contact with the criminal justice system

A higher proportion of the fringe or radical content group (3.7% vs 2.8%; $F$=6.3, $p<0.05$) and the overlap group (10.5% vs 2.8%; $F$=56.8, $p<0.01$) reported having been arrested, charged, summonsed, convicted or found guilty of an offence (ie contact with the criminal justice system) as a juvenile compared with respondents who had seen neither type of content. The overlap group were more than twice as likely as the CSAM group (10.5 % vs 4.6%; $F$=4.2, $p<0.05$) and almost three times as likely as the fringe or radical content group (10.5% vs 3.7%; $F$=33.1, $p<0.01$) to have had contact with the criminal justice system as a juvenile. Similarly, the proportions of the CSAM group (11.1%; $F$=4.4, $p<0.05$), fringe or radical content group (8.9%; $F$=18.6, $p<0.01$) and overlap group (17.8%; $F$=59.3, $p<0.01$) who had contact with the criminal justice system as an adult were significantly higher than for the group who had viewed neither type of content (6.5%). The overlap group had the highest proportion of respondents with an adult criminal history—double that of the fringe or radical content group (17.8% vs 8.9%; $F$=27.8, $p<0.01$).

**Table 4: Contact with the criminal justice system as a juvenile or adult, by group (%)**

| | Viewed neither CSAM nor fringe or radical content (*n*=5,822) | Viewed only CSAM (*n*=150) | Viewed only fringe or radical content (*n*=4,768) | Viewed both CSAM and fringe or radical content (*n*=387) |
|---|---|---|---|---|
| Juvenile delinquency | 2.8 | 4.6 | 3.7** | 10.5**##xx |
| Adult offending | 6.5 | 11.1* | 8.9** | 17.8**xx |

For comparisons with respondents who did not view either CSAM or fringe or radical content, ***p*<0.01 and **p*<0.05
For comparisons with respondents who only viewed CSAM, ##*p*<0.01 and #*p*<0.05
For comparisons with respondents who only viewed fringe or radical content, xx*p*<0.01 and x*p*<0.05
Source: Survey of social and political attitudes in Australia 2022 [computer file]

## Self-reported diagnosed mental health disorders

Respondents were asked whether they had ever been diagnosed with a mental health disorder. The overlap group had the highest rate of diagnosed mental disorders overall (58.5%; *F*=101.1, *p*<0.01), followed by the CSAM group (47.9%; *F*=15.4, *p*<0.01) and then the fringe or radical content group (39.4%; *F*=58.9, *p*<0.01). This pattern held for all types of disorders other than psychotic disorders.

Among the overlap group, 58.5 percent reported they had been diagnosed with a mental health disorder, compared to 47.9 percent of the CSAM group (*F*=4.4, *p*<0.05), 39.4 percent of the fringe or radical content group (*F*=47.5, *p*<0.01) and 31.8 percent of those who viewed neither type of content (*F*=101.1, *p*<0.01). Of note, the rate of diagnosed psychotic disorders in the overlap group was six times greater than in the CSAM group (6.8% vs 0.9%; *F*=5.6, *p*<0.01), and four times greater than in the fringe or radical content group (6.8% vs 1.7%; *F*=32.9, *p*<0.01).

**Table 5: Self-reported diagnosed mental health disorder, by group**

| | Viewed neither CSAM nor fringe or radical content (*n*=5,822) | Viewed only CSAM (*n*=150) | Viewed only fringe or radical content (*n*=4,768) | Viewed both CSAM and fringe or radical content (*n*=387) |
|---|---|---|---|---|
| Any diagnosed mental health disorder | 31.8 | 47.9** | 39.4**# | 58.5**#xx |
| Depressive disorder | 19.8 | 35.5** | 24.6**# | 35.3**xx |
| Anxiety disorder | 21.3 | 32.4* | 27.0** | 38.4**xx |
| Psychotic disorder | 1.5 | 0.9 | 1.7 | 6.8**##xx |
| Personality disorder | 1.9 | 6.1** | 2.7*# | 8.6**xx |
| Substance use disorder | 3.5 | 6.9* | 4.9** | 9.9**xx |
| Other psychiatric condition | 4.5 | 14.5** | 6.2**## | 13.8**xx |

For comparisons with respondents who did not view either CSAM or fringe or radical content, ***p*<0.01 and **p*<0.05
For comparisons with respondents who only viewed CSAM, ##*p*<0.01 and #*p*<0.05
For comparisons with respondents who only viewed fringe or radical content, xx*p*<0.01 and x*p*<0.05
Source: Survey of social and political attitudes in Australia 2022 [computer file]

## Discussion

This research has shown that there is a considerable overlap between people who view CSAM and who view fringe and radical content online. In particular, the majority of people who view CSAM also view content they described as fringe or radical. The profile of this overlap group also differed from that of respondents who viewed either one or neither type of content. Respondents who viewed both types of content were more likely to be younger and male, to have had contact with the criminal justice system as juveniles and adults and to have been diagnosed with certain mental illnesses. There were also important differences in their online behaviours and the platforms they used.

Conversely, respondents who viewed only CSAM or only fringe or radical content were relatively similar to one another. Respondents who viewed fringe or radical content tended to be older, used the internet less often for work and personal reasons, were less likely to use TikTok and were less likely to report having ever been diagnosed with mental health disorders such as depressive, anxiety, personality or substance use disorders.

The implications of viewing these types of harmful online content extend beyond the consumption of the content alone. Cubitt and Morgan (2024) found that, among those who had seen fringe or radical content, 29.1 percent had encountered a group promoting extreme, fringe or radical views online, 6.9 percent reported becoming a member of a group, and 12.7 percent reported sharing fringe or radical content with at least one other person, commonly on mainstream social media platforms. Further, being a member of a group promoting fringe or radical content was associated with increased sharing of that content with other internet users. Respondents who view fringe and radical content online and share that as part of an online group may, given the overlap observed here, be facilitators of other types of content, including CSAM.

### Implications for prevention

Most research and intervention approaches relating to either online CSAM or fringe or radical content do not consider co-occurring access to other types of content. This research has shown that, within a 12-month period, 65 percent of respondents who viewed CSAM had also viewed fringe or radical content, and seven percent of those who accessed fringe or radical content also viewed CSAM. Whether respondents had intentionally or unintentionally viewed one or both types of content was unclear—it could be that where one type of content was sought, the other was also present. However, it could also be that a large proportion of the sample were seeking both types of content, particularly those who were looking for CSAM online. Regardless, these findings have important implications for intervention approaches seeking to address access to problematic online content.

Prevention activities designed to disrupt access to online content tend to focus on the platform hosting the content or the format of the content itself. For example, common approaches to combat access to online CSAM include intervention messaging, content removal, peer-to-peer monitoring and messaging campaigns (Hunn et al. 2023). These approaches are similar to those designed to prevent access to fringe or radical content (Wolbers et al. 2023). While the tactics for prevention and disruption have been similar, the platforms or internet users targeted have differed depending on the content. This research has shown that the people who viewed CSAM and who viewed fringe or radical content were, in fact, relatively similar. The profile of respondents who viewed both types of content was different, and this may help guide efforts to target this high-risk group.

A central assumption underpinning prevention approaches is that individuals who access CSAM only access CSAM, and individuals who access fringe or radical content only access fringe or radical content. Disrupting access to specific types of content is therefore thought to reduce future access and downstream offending. However, the presence of a group who view both CSAM and fringe or radical content online, and who are meaningfully different from those who view only one type of content, is important. Prevention strategies may be more effective and reap wider benefits when they target both types of harmful content, the platforms that host both types or the internet users who seek them.

Separately, it is also possible that the treatment needs of individuals who access either CSAM or fringe or radical content may be different to those who access both. For example, our findings suggest those who accessed both types of content may experience more diagnosed psychiatric conditions and a more diverse range of them. Treatment interventions for CSAM consumers are different to interventions for those who radicalise, and those who access both types of content may need different treatment again. Further research should be undertaken to inform psychosocial intervention and prevention approaches for those who access multiple types of harmful content.

### Directions for future research

These findings also have implications for the insights that can be gained by using different methodologies to research online environments. For example, emerging research into online fringe or radical content has focused on platforms known to contain extreme content, such as Telegram (Vergani et al. 2022), Twitter (now X; Froio & Ganesh 2019) or Gab (Jasser et al. 2023). The same is often true of research on access to CSAM, which has focused on platforms such as Omegle (Salter & Sokolov 2024) or the darknet (Ngo et al. 2024). While these studies are important to our understanding of these online environments, the characteristics and behaviours they analyse are specific to those platforms (including the particular users who post or actively engage on them). This study suggests that engagement with online platforms hosting CSAM or fringe or radical content may be more complex. Individuals who view this content may not specialise in the platform they use to access it or seek only one type of content. Future research should aim to take a broader focus and not view these types of content, internet users and platforms in isolation from one another.

## References

*URLs correct as at September 2024*

Argentino MA, Barrett G & Tyler MB 2024. *764: The intersection of terrorism, violent extremism, and child sexual exploitation.* Global Network on Extremism & Technology: Washington. https://gnet-research.org/2024/01/19/764-the-intersection-of-terrorism-violent-extremism-and-child-sexual-exploitation/

Australian Bureau of Statistics (ABS) 2023. National, state and territory population, September 2022. https://www.abs.gov.au/statistics/people/population/national-state-and-territorypopulation/sep-2022

Australian Security Intelligence Organisation 2023. Director-General's Annual Threat Assessment 2023. https://www.asio.gov.au/director-generals-annual-threat-assessment-2023

Balfe M, Gallagher B, Masson H, Balfe S, Brugha R & Hackett S 2015. Internet child sex offenders' concerns about online security and their use of identity protection technologies: A review. *Child Abuse Review* 24(6): 427–439. https://doi.org/10.1002/car.2308

Broadhurst R 2019. Child sex abuse images and exploitation materials. In R Leukfeldt & T Holt (eds), *Cybercrime: The human factor*. Routledge: 310–336. https://doi.org/10.4324/9780429460593-14

Broadhurst R, Woodford-Smith H, Maxim D, Sabol B, Orlando S, Chapman-Smith B & Alazab M 2017. *Cyber terrorism: Research review*. Canberra: Australian National University, Cybercrime Observatory. https://regnet.anu.edu.au/research/publications/6961/cyber-terrorism-research-review-research-report-australian-national

Brown R 2023. Prevalence of viewing online child sexual abuse material among Australian adults. *Trends & issues in crime and criminal justice* no. 682. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77239

Brown R 2022. *Eliminating online child sexual abuse material*. Abingdon: Routledge. https://doi.org/10.4324/9781003327264

Bursztein E, Clarke E, DeLaune M, Eliff DM, Hsu N, Olson L, Shehan J, Thakur M, Thomas K & Bright T 2019. Rethinking the detection of child sexual abuse imagery on the internet. Paper presented at the World Wide Web Conference, San Francisco, CA: 2601–2607

Cubitt T & Morgan A 2024. Exposure to and sharing of fringe or radical content online. *Trends & issues in crime and criminal justice* no. 705. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77734

Droogan J & Waldek L 2018. Should we be afraid of cyber-terrorism? *International Journal of Electronic Security and Digital Forensics* 10(3): 242–254. https://doi.org/10.1504/IJESDF.2018.093017

Federal Bureau of Investigation 2023. Violent online groups extort minors to self-harm and produce child sexual abuse material. Public Service Announcement Alert Number I-091223-PSA. https://www.ic3.gov/Media/Y2023/PSA230912

Froio C & Ganesh B 2019. The transnationalisation of far right discourse on Twitter. *European Societies* 21(4): 513–539. https://doi.org/10.1080/14616696.2018.1494295

Gill P, Corner E, Conway M, Thornton A, Bloom M & Horgan J 2017. Terrorist use of the internet by the numbers: Quantifying behaviors, patterns, and processes. *Criminology & Public Policy* 16(1): 99–117. https://doi.org/10.1111/1745-9133.12249

Hamid N & Ariza C 2022. *Offline versus online radicalization: Which is the bigger threat? Tracing outcomes of 439 jihadist terrorists between 2014–2021 in 8 Western countries*. Global Network on Extremism and Technology. https://gnet-research.org/2022/02/21/offline-versus-online-radicalisation-which-is-the-bigger-threat/

Hassan G et al. 2018. Exposure to extremist online content could lead to violent radicalization: A systematic review of empirical evidence. *International Journal of Developmental Science* 12(1–2): 71–88. https://doi.org/10.3233/DEV-170233

Holt TJ, Cale J, Leclerc B & Drew J 2020. Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses. *Aggression and Violent Behavior* 55: 101464. https://doi.org/10.1016/j.avb.2020.101464

Hunn C, Watters P, Prichard J, Wortley R, Scanlan J, Spiranovic C & Krone T 2023. How to implement online warnings to prevent the use of child sexual abuse material. *Trends & issues in crime and criminal justice* no. 669. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78894

Jasser C, McSwiney J, Pertwee E & Zannettou S 2023. 'Welcome to the #GabFam': Far-right virtual community on Gab. *New Media & Society* 25(7): 1728–1745. https://doi.org/10.1177/14614448211024546

Johnson B & Patel P 2019. Multi-million pound funding to protect child abuse victims and track down offenders. https://www.gov.uk/government/news/multi-million-pound-funding-to-protect-child-abuse-victims-and-track-down-offenders

Kenyon J, Binder JF & Baker-Beall C 2023. Online radicalization: Profile and risk analysis of individuals convicted of extremist offences. *Legal and Criminological Psychology* 28(1): 74–90. https://doi.org/10.1111/lcrp.12218

Morgan S & Lambie I 2019. Understanding men who access sexualised images of children: Exploratory interviews with offenders. *Journal of Sexual Aggression* 25(1): 60–73. https://doi.org/10.1080/13552600.2018.1551502

Ngo VM, Gajula R, Thorpe C & Mckeever S 2024. Discovering child abuse material creators' behaviors and preferences on the dark web. *Child Abuse & Neglect* 147: 106558. https://doi.org/10.1016/j.chiabu.2023.106558

Nienierza A, Reinemann C, Fawzi N, Riesmeyer C & Neumann K. 2021. Too dark to see? Explaining adolescents' contact with online extremism and their ability to recognize it. *Information, Communication & Society* 24(9): 1229–1246. https://doi.org/10.1080/1369118X.2019.1697339

Nuraniyah N 2018. Not just brainwashed: Understanding the radicalization of Indonesian female supporters of the Islamic State. *Terrorism and Political Violence* 30(6): 890–910. https://doi.org/10.1080/09546553.2018.1481269

Rao JNK & Scott AJ 1984. On chi-squared tests for multiway tables with cell proportions estimated from survey data. *Annals of Statistics* 12: 46–60. https://doi.org/10.1214/aos/1176346391

Rao JNK & Scott AJ 1981. The analysis of categorical data from complex sample surveys: Chi-squared tests for goodness-of-fit and independence in two-way tables. *Journal of the American Statistical Association* 76: 221–230. https://doi.org/10.1080/01621459.1981.10477633

Salter M & Sokolov S 2024. "Talk to strangers!" Omegle and the political economy of technologically-facilitated child sexual exploitation. *Journal of Criminology* 57(1): 121–137. https://doi.org/10.1177/26338076231194451

Salter M, Woodlock D, Whitten T, Tyler M, Naldrett G, Breckenridge J, Nolan J, Peleg N 2023. *Identifying and understanding child sexual offending behaviours and attitudes among Australian men*. Sydney: University of New South Wales. https://www.humanrights.unsw.edu.au/news/worlds-largest-child-sexual-abuse-perpetration-prevalence-study-recommends-significant-investment-early-intervention-measures

Seto MC & Ahmed AG 2014. Treatment and management of child pornography use. *Psychiatric Clinics of North America* 37(2): 207–214. https://doi.org/10.1016/j.psc.2014.03.004

Vergani M, Arranz AM, Scrivens R & Orellana L 2022. Hate speech in a Telegram conspiracy channel during the first year of the COVID-19 pandemic. *Social Media + Society* 8(4). https://doi.org/10.1177/20563051221138758

WePROTECT Global Alliance 2021. *Global threat assessment 2021: Working together to end the sexual abuse of children online*. https://www.weprotect.org/global-threat-assessment-21/#report

Westlake B 2020. The past, present, and future of online child sexual exploitation: Summarizing the evolution of production, distribution, and detection. In T Holt & A Bossler (eds), *The Palgrave handbook of international cybercrime and cyberdeviance*. Cham: Palgrave Macmillan: 1225–1253. https://doi.org/10.1007/978-3-319-78440-3_52

Wolak J, Liberatore M & Levine BN 2014. Measuring a year of child pornography trafficking by US computers on a peer-to-peer network. *Child Abuse & Neglect* 38(2): 347–356. https://doi.org/10.1016/j.chiabu.2013.10.018

Wolbers H, Dowling C, Cubitt T & Kuhn C 2023. Understanding and preventing internet facilitated radicalisation. *Trends & issues in crime and criminal justice* no. 637. Canberra: Australian institute of Criminology. https://doi.org/10.52922/ti77024

Wolfowicz M, Hasisi B & Weisburd D 2022. What are the effects of different elements of media on radicalization outcomes? A systematic review. *Campbell Systematic Reviews* 18(2): e1244. https://doi.org/10.1002/cl2.1244

Wortley R & Smallbone S 2012. *Internet child pornography: Causes, investigation, and prevention*. Global crime and justice series. Santa Barbara, CA: Praeger

This work was originally published as *Trends & issues in crime and criminal justice* no. 708.
It is available online at **https://doi.org/10.52922/ti77710**.

# 7. Artificial intelligence and child sexual abuse: A rapid evidence assessment

Heather Wolbers, Timothy Cubitt and Michael John Cahill

## Introduction

In recent years, the development of artificial intelligence (AI) has rapidly increased, with availability of this technology significantly expanding. Development of AI technologies has extended to the field of child sexual abuse (CSA), with the scope and scale of the problem—particularly online—becoming too great for manual human-led approaches to manage effectively. However, the use of AI in this field has extended beyond prevention. In early 2023, the US National Center for Missing and Exploited Children received reports of 'fake' child sexual abuse materials (CSAM) that offenders had produced with the assistance of generative AI tools (Murphy 2023). Similarly, Australia's eSafety Commissioner has noted reports of children using AI to generate sexually explicit images of other children, suggesting that it was an indication of a more widespread issue (Long 2023).

This review considers the current state of research literature studying the use of AI in the field of CSA, focusing on studies investigating the use of AI for offending and the prevention and disruption of CSA.

### Artificial intelligence for child sexual abuse offending

According to media reports, surveys and academic reviews, AI technologies are increasingly playing diverse roles in the creation of CSAM, including fabrication (eg CSAM deepfakes), 'nudifying' pictures of clothed children, and manipulating images or videos to depict known or unknown children in sexually abusive scenarios (Milmo 2023; Okolie 2023). Existing CSAM has been used to train AI models, meaning offenders are using AI to produce new depictions of previously abused children. Further, reports indicate that offenders are using AI to alter photos from victims' social media and other online posts and using these altered images to sexually extort the victims (Garriss & DeMarco 2023). According to a survey of 1,040 people aged nine to 17 years in the United States, one in 10 (11%) minors said they knew of cases where their peers had used AI to create sexually explicit images of other minors (Thorn 2024b).

The Internet Watch Foundation found that, in a single month, 20,254 AI-generated images were posted to a CSAM forum on the darknet (Internet Watch Foundation 2023). Concerns have been raised that the ability to generate CSAM using AI could support an increase in CSAM consumption. Growth in AI-generated CSAM creates significant challenges for law enforcement, who work to detect and prevent the distribution of CSAM online. Ultimately, increases in the volume of CSAM online may influence the ability to investigate CSA, as AI-generated can be indistinguishable from real CSAM (Theil, Stroebel & Portnoff 2023). The malicious use of AI technologies for the production of CSAM is growing and is likely to continue to grow without multi-sector intervention (Theil, Stroebel & Portnoff 2023).

## Artificial intelligence for the prevention and disruption of child sexual abuse

As identified in academic research, there are a diverse range of cyber strategies used to combat online CSA (Edwards et al. 2021; Singh & Nambiar 2024). As the field of AI continues to develop, so too does the development of CSA disruption strategies that use AI. For example, published research has shown that AI could be used to identify suspicious financial transactions procuring CSA (eg Cubitt, Napier & Brown 2021; Henseler & de Wolf 2019) or aid in law enforcement investigations by examining CSAM (eg Brewer et al. 2023; Dalins et al. 2018; Westlake et al. 2022). AI technologies may ease the burden on law enforcement by reducing the risk of psychological harms among CSA investigators (Puentes et al. 2023), while increasing the capacity and timeliness of investigations. Further, AI technologies can have a much larger reach across online spaces than traditional methods of prevention and disruption.

While CSAM can be detected and removed across online spaces with hashes (ie unique digital fingerprints), this method is limited to known CSAM on platforms proactively using hashes, meaning there is limited efficacy and it cannot stop the upload and proliferation of undetected, new or edited content. AI has the potential to help address this challenge. For instance, Thorn has developed a machine learning tool to automatically detect, review and report CSAM at scale (Thorn 2024a). This tool is used to screen new content that gets uploaded to Flickr and other online platforms—a task too large for human moderation alone. Beyond detecting CSAM, AI has a range of potential uses for addressing CSA. These include conversation analysis, chatbots, honeypots and web crawlers, all of which show promise in combating CSA, albeit with some significant limitations (eg narrow scope or generalisability, privacy and legal concerns, and lack of robust evaluation; Singh & Nambiar 2024).

## Study aims

AI is the ability of a computer system to simulate human intelligence, such as learning, problem solving, reasoning and decision making—all with some level of autonomy (High-Level Expert Group on Artificial Intelligence 2019). There are several domains of AI, including machine learning (in which an algorithm is trained to learn patterns in existing data), computer vision (interpreting visual information), natural language processing (understanding and generating human language), and generative AI (creating original content). We considered the domains of AI to develop search criteria and identify literature examining AI and its intersection with CSA.

This study aims to establish current AI capabilities in relation to CSA, with the intention of identifying key areas of progress and gaps where further research is required. A rapid evidence assessment was conducted to address the following research questions:

- What are the uses of AI as a part of CSA offending and what are the areas of future risk?
- What are the uses of AI as a part of CSA prevention and disruption, and what are the areas of future potential?
- What are the key gaps in current research that should be addressed?

# Method

## Search strategy

Rapid evidence assessments are accelerated systematic reviews of research undertaken in a restrictive time frame. This study draws on research published in English between January 2010 and March 2024. Studies were excluded if they did not discuss AI in the context of CSA, include primary data (ie reviews or conceptual studies), explain the study's methodology in sufficient detail (eg if they did not detail the data source, the sample or data management for analysis), or have a direct application to CSA. We excluded studies examining the use of AI in medical settings to detect CSA, as a systematic review was recently published on this topic, which identified seven studies that examined the use of AI for predicting child abuse and neglect using medical or protective service data (Lupariello et al. 2023). Of note, our search yielded seven studies from medical settings, all of which were excluded for secondary reasons (ie they did not focus on detection or prevention of CSA).

The Australian Institute of Criminology's JV Barry Library searched 13 databases and relevant websites: the JV Barry Library catalogue, the Australian Institute of Criminology, EBSCO, ProQuest Criminal Justice, DeepDyve, arXiv.org, IEEE Xplore digital library, Office of the eSafety Commissioner, International Centre for Missing and Exploited Children, National Center for Missing and Exploited Children, Australian Centre to Counter Child Exploitation, Thorn and Google Scholar.

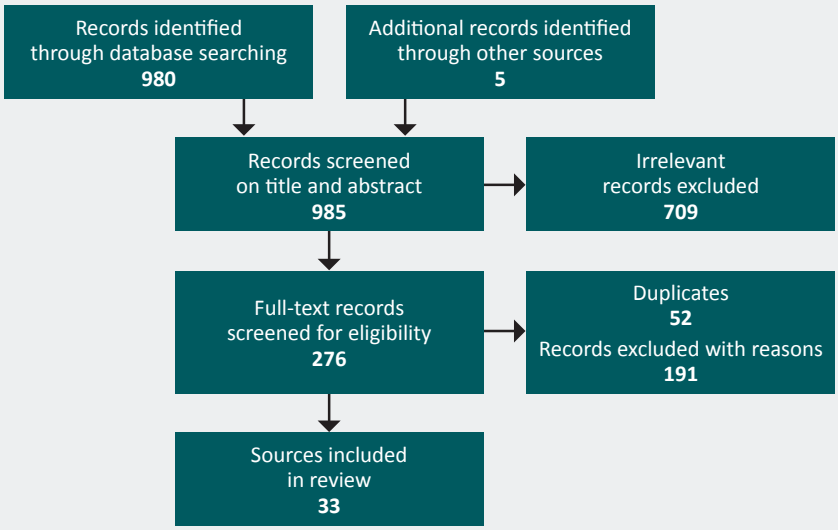
The search terms used combined keywords from three categories capturing AI and its relevant sub-domains, child sexual abuse, and specifying the group of concern to be individuals under 18 years of age:

- *Artificial intelligence* ("machine learning" OR "artificial intelligence" OR algorithm* OR "deep learning" OR "unsupervised learning" OR "reinforcement learning" OR "generative artificial intelligence" OR "natural language processing" OR "computer vision" OR chatbot OR "image classification" OR "object detection" OR "augmented reality" OR "big data" OR "neural network");
- *Child sexual abuse* ("child sexual abuse" OR CSAM OR CSEM OR "child abuse material" OR "live streaming" OR "child exploitation" OR "child sexual abuse material" OR "child exploitation material" OR "image-based sexual abuse" OR "image-based abuse" OR "technology-facilitated sexual violence" OR online "sexual exploitation of children" OR "child pornography" OR "indecent images of children" OR grooming); and
- *Child* (child OR children OR "young person" OR "young people" OR adolescent OR teenage* OR youth OR minor OR "young adult").

## Screening process

The rapid evidence search yielded 980 records. Five additional studies were identified through wider reading, resulting in identification of an initial 985 records. Titles and abstracts were screened to exclude irrelevant studies (*n*=709). This screening yielded 276 records, of which 52 were identified as duplicates and removed. The remaining 224 sources were assessed for eligibility against the selection criteria with full-text screening, and 191 were excluded because they did not meet the selection criteria. In total, the search yielded 33 sources providing primary information on the role of AI in relation to CSA (see Figure 1).

**Figure 1: Literature screening process**




### Limitations

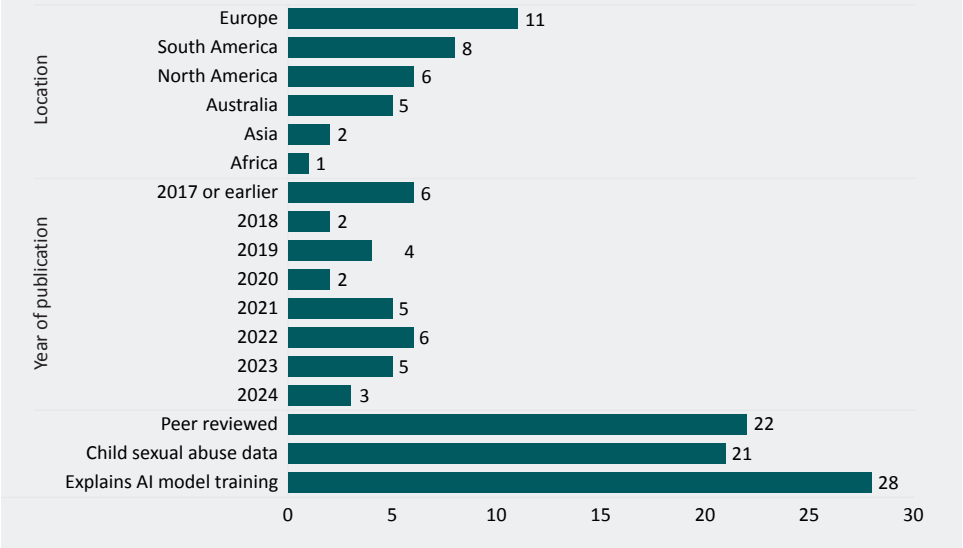
Rapid evidence assessments do not provide the same exhaustive depth or detail as a full systematic review (Ganann, Ciliska & Thomas 2010). Databases that yielded a large number of hits were not searched in full. Rather, the first several hundred items returned by the search were screened, meaning that the most relevant sources were captured. However, search results were not screened exhaustively. Given that search results were presented in order of relevance to the search terms used, the number of sources missed by using this methodology is likely limited.

## Results

### Study characteristics

Research at the intersection of AI and CSA was identified from Europe (*n*=11), South America (*n*=8), North America (*n*=6), Australia (*n*=5), Asia (*n*=2) and Africa (*n*=1). Identified research was published between 2011 to 2024, with a notable increase from 2020 onward. Two-thirds were peer reviewed (*n*=22). Non-peer reviewed sources included conference papers (*n*=9) or pre-prints (*n*=2). While research primarily relied on data from cases of CSA (ie CSAM files, or information on convicted CSA offenders), five studies used data on suspected rather than proven child sexual exploitation (eg reports to a hotline, risky online conversations). Five studies relied on interactions between suspected or known offenders and adults posing as children, and two relied on peripheral datasets using semi-nude non-sexual images of children and a database of faces for age estimation. The majority of included studies explained how their AI model was trained (*n*=28), with five using unsupervised modelling approaches. All identified research examined uses of AI for the prevention, disruption, detection or investigation of CSA, with no studies examining the uses of AI to perpetrate CSA. As reports of malicious use of AI to perpetrate CSA only began to emerge recently (Long 2023; Murphy 2023), it seems unlikely that enough time has passed for empirical research to have been produced on this subject matter.

**Figure 2: Characteristics of included studies (*n*=33)**




## Quality of evidence regarding efficacy of artificial intelligence in the context of child sexual abuse

Among the 33 reviewed studies, 28 attempted to evaluate the discussed AI tool—typically considering accuracy, precision, recall or another metric specific to the intended goal. When a tool was evaluated, the employed metrics and study aims were diverse, meaning cross-study comparisons of efficacy were not possible. Additionally, the data or sample used for testing—particularly when the data were synthetically produced, or relied on a small or non-generalisable sample—raised questions about how well the model would translate to a real-world setting. Only a minority of studies tested the tool in a real-world or near real-world setting (Brewer et al. 2023; Dalins et al. 2018; Guerra & Westlake 2021; Jin et al. 2024; Ngo et al. 2024; Peersman et al. 2016; Westlake et al. 2022). Due to these limitations, any findings regarding efficacy for each included study should be interpreted with some caution. For example, while a tool may demonstrate a high level of efficacy, this does not necessarily mean the tool would perform well if applied in the real world.

## Artificial intelligence and child sexual abuse

Table 1 presents a summary of the AI tools discussed in published literature. These included tools used for detection, examination or investigation of CSAM or child sexual offenders online.

**Table 1: Summary of how artificial intelligence is used in the context of child sexual abuse in the studies reviewed**

| Intention of the tool | Approach used |
|---|---|
| Detect CSAM | Detect CSAM on personal computers using file names and file paths (2, 22, 23) |
| | Detect CSAM using a combination of tools such as pornography detection, age estimation, skin tone/nudity identifier (8, 9, 16, 21, 22, 24, 25, 29, 32) |
| Aid in the investigation of CSAM | Separate CSAM into discrete categories by type (8, 15) |
| | Determine the age of children in CSAM (11) |
| | Match victims and offenders across CSAM videos using facial and voice recognition (4, 33) |
| | Identify patterns in the locations and folder or file naming practices of websites with CSAM (12) |
| Detect online child sexual offenders | Analyse the language used in online conversations to identify threats to children (1, 3, 5, 13, 17, 26, 30) |
| | Analyse conversations between children and offenders to identify whether offenders intend to contact offend or not (31) |
| | Distinguish real children from adults pretending to be children in chat rooms (17) |
| | Using a chatbot to interact with suspects and profile their interest in CSAM (18, 27) |
| | Scrape hashtags and images from tweets in real time to detect suspected human trafficking of minors (10) |
| Aid in understanding online child sexual offenders | Crawl the darknet to collect data on the behaviours of child sexual offenders who access and participate on dark websites (14) |
| | Analyse posts about CSAM and associated metadata to understand the characteristics, behaviours and motivations of CSAM creators (20) |
| | Understand the characteristics and typologies of offenders who live stream CSA (6, 7) |
| | Detect other files shared online by individuals who have shared known CSAM files (22) |
| Other | Analyse text-based reports of child sexual abuse (25) |

Note: 1—Agarwal et al. 2022; 2—Al-Nabki et al. 2023; 3—Anderson et al. 2019; 4—Brewer et al. 2023; 5—Cardei & Rebedea 2017; 6—Cubitt, Napier & Brown 2021; 7—Cubitt, Napier & Brown 2023; 8—Dalins et al. 2018; 9—Gangwar et al. 2021; 10—Granizo et al. 2020; 11—Grubl & Lallie 2022; 12—Guerra & Westlake 2021; 13—Isaza et al. 2022; 14—Jin et al. 2024; 15—Laranjeira et al. 2022; 16—Macedo, Costa & dos Santos 2018; 17—Meyer 2015; 18—Murcia Triviño et al. 2019; 19—Ngejane et al. 2021; 20—Ngo et al. 2024; 21—Oronowicz-Jaśkowiak et al. 2024; 22—Peersman et al. 2016; 23—Pereira et al. 2021; 24—Polastro & Eleuterio 2012; 25—Puentes et al. 2023; 26—Razi et al. 2023; 27—Rodríguez et al. 2020; 28—Rondeau et al. 2022; 29—Sae-Bae et al. 2014; 30—Seedall, MacFarlane & Holmes 2019; 31—Seigfried-Spellar et al. 2019; 32—Ulges & Stahl 2011; 33—Westlake et al. 2022

### *Artificial intelligence for detecting and investigating child sexual abuse material*

The most common use of AI in the identified research was to detect CSAM. The primary intention of the AI tools studied was to reduce the burden of manual processing of CSAM by investigators, thereby mitigating mental health impacts, while reducing the time needed to identify CSAM among very large datasets. Tools designed to detect CSAM tended to combine age identifiers (Dalins et al. 2018; Gangwar et al. 2021; Macedo, Costa & dos Santos 2018; Rondeau et al. 2022; Sae-Bae et al. 2014), with skin tone, nudity or pornography detectors (Dalins et al. 2018; Gangwar et al. 2021; Laranjeira et al. 2022; Macedo, Costa & dos Santos 2018; Oronowicz-Jaśkowiak et al. 2024; Peersman et al. 2016; Polastro & Eleuterio 2012; Rondeau et al. 2022; Sae-Bae et al. 2014).

Other models detected CSAM by analysing the words used to describe the picture (Peersman et al. 2016; Ulges & Stahl 2011) or language embedded in audio (Peersman et al. 2016). Three studies discussed tools that analyse file names or file paths to estimate the likelihood of them containing CSAM (Al-Nabki et al. 2023; Peersman et al. 2016; Pereira et al. 2021). Importantly, tools were frequently designed to be implemented in a specific setting, such as on peer-to-peer networking websites (eg Peersman et al. 2016) or on a personal computer or device (eg Polastro & Eleuterio 2012). Many of these studies used authentic CSA data sources and reported that the tool of focus performed well at completing the intended task (Al-Nabki et al. 2023; Gangwar et al. 2021; Oronowicz-Jaśkowiak et al. 2024; Peersman et al. 2016; Pereira et al. 2021; Polastro & Eleuterio 2012). However, others showed limited efficacy (Dalins et al. 2018; Puentes et al. 2023; Ulges & Stahl 2011).

Notably, Peersman and colleagues (2016) described a toolkit that performed multiple functions. While designed primarily to detect CSAM on peer-to-peer networks by analysing filenames, images and audio, the tool also flagged files shared by individuals who have shared known CSAM. This tool was evaluated using real-world CSA case data, demonstrating usefulness in law enforcement settings and considerable accuracy in detecting CSAM when combining filename and image classification.

Beyond the detection of CSAM, several studies discussed uses of AI to assist investigations in other ways. One study aimed to categorise the content of CSAM to aid with triaging (eg solo, non-penetrative, penetrative; Dalins et al. 2018), while another extracted features and labels of CSAM to describe the content without it ever being viewed (Laranjeira et al. 2022). Another important use was to extract and match the biometric features of victims and perpetrators shown in CSAM, allowing for a rapid detection of media associated with an investigation and the identification of links between files (Brewer et al. 2023; Westlake et al. 2022). Finally, a web crawler was designed to find patterns in the locations and CSAM naming conventions of websites with known CSAM (Guerra & Westlake 2021). Three of these studies tested the performance of the AI tool (Dalins et al. 2018; Laranjeira et al. 2022; Westlake et al. 2022), and just one demonstrated strong results. Specifically, Westlake and colleagues (2022) were able to identify and match victims and offenders across a test sample of authentic CSAM files with a high true match rate (between 93.8% and 98.8%) and a low false match rate (between 1.0% and 5.0%).

### *Artificial intelligence for detecting and understanding child sexual abuse offenders*

Two studies described a tool designed to initiate and hold conversations with potential online CSA offenders (ie a chatbot; Murcia Triviño et al. 2019; Rodríguez et al. 2020). This chatbot used generative and rule-based models to produce conversational posts and replies. Based on these interactions, the tool then described each suspect's behaviour, classifying their disposition towards online child sexual offending as indifferent, interested or perverted. The efficacy of this classification was not numerically measured.

Several studies described AI methods designed to understand child sexual offenders through their online behaviours. A tool produced by Granizo and colleagues (2020) scraped posts from X (formerly Twitter) in real time to identify suspected cases of child trafficking. This tool demonstrated some efficacy at recognising the gender and age of individuals depicted in images by analysing their faces or torso.

Two studies discussed tools operating on the darknet. The first study discussed a web crawler that collected data on darknet ecosystems, detecting relevant content and providing information designed to reduce the anonymity of perpetrators, such as links to the surface web that may be used to trace darknet operators (Jin et al. 2024). Similarly, a tool developed by Ngo and colleagues (2024) processed CSAM discussions on the darknet and provided insights into the characteristics, behaviours and motivation of CSAM creators. Both tools performed well in identifying the content of interest and were able to reveal meaningful information about online offending environments and offenders.

Peersman and colleagues (2016) described a model to detect other online files shared by those known to distribute CSAM online, while two studies used machine learning to analyse the characteristics and offending history of individuals who live streamed CSA (Cubitt, Napier & Brown 2023, 2021). In the 2021 study, the model had notable success in identifying individuals who would engage in prolific live streaming of CSA, successfully classifying more than 80 percent of cases (AUROC=0.85; Cubitt, Napier & Brown 2021).

### *Other uses of artificial intelligence in the context of child sexual abuse*

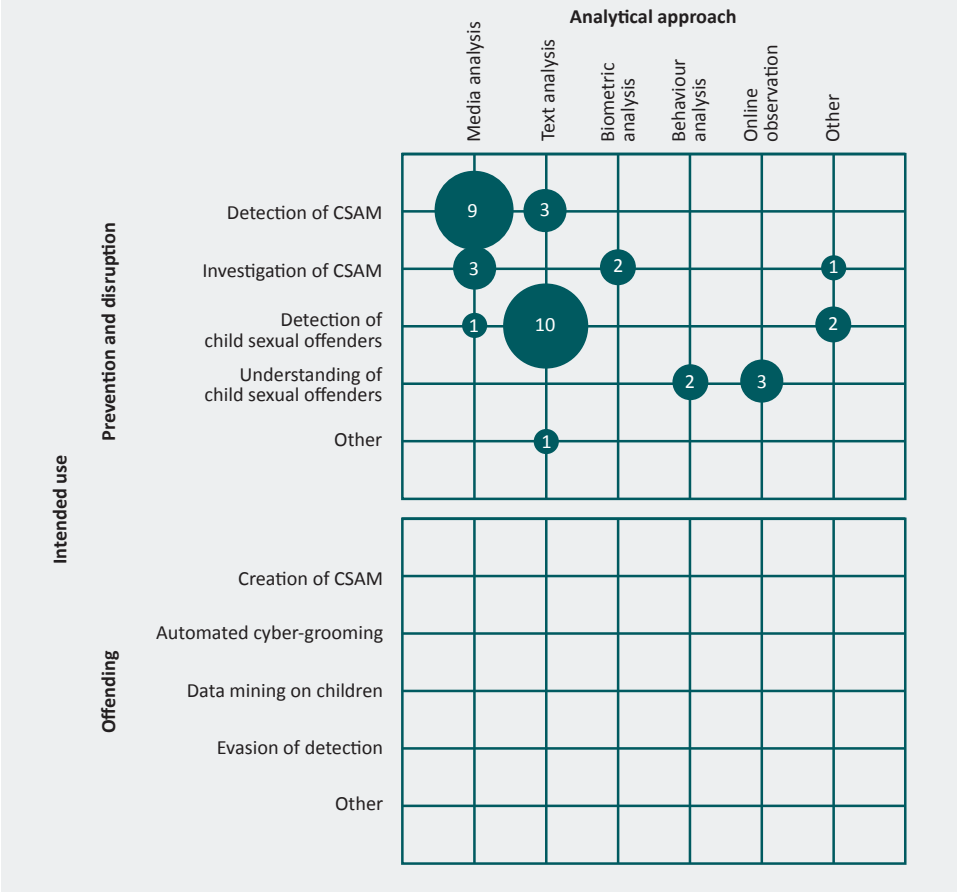
One final tool used a large language model to identify the subject, degree of criminality, and level of impact relating to reports of CSA to a hotline (Puentes et al. 2023). Using this method, the authors aimed to automate the analysis of CSA reports, consequently expediting the process while reducing the exposure of analysts to potentially harmful content. The authors concluded that the approach was an appropriate starting point, but the efficacy was limited.

## Evidence and gap map

We constructed an evidence and gap map plotting the most common analytical domains of AI that were reported, and how they were used in the context of CSA (Figure 3). Importantly, these do not represent all possible AI capabilities and uses in the context of CSA. The analytical procedure and intended use categories broadly capture the most common AI capabilities and uses outlined in the research, but the figure does not critique the quality of evidence or efficacy of the AI technologies discussed. Dots on the graph show where the analytical approach and the intended uses intersect, with the size of the dot reflecting the number of studies. Intersections without dots indicate an absence of evidence, highlighting areas requiring further research.

The evidence and gap map highlighted a significant gap in evidence relating to how AI is used in the perpetration of child sexual offending, with no literature returned in this search. The two most developed areas of research focused on analysis of media (images, audio and videos) to detect CSAM, and text analysis to detect child sexual offenders.

**Figure 3: Evidence and gap map for research on the uses of artificial intelligence in the context of child sexual abuse**




## Discussion

At the time of this review, research has examined AI tools used to detect CSAM or CSA offenders, to aid with investigations, and to improve our understanding of online environments where CSA is produced and shared. The two most common uses of AI were analysing images, audio or video to detect CSAM without requiring humans to view the content, and language processing to detect child sexual offenders through online conversations. Of note, several of the tools described undertook more than one task. For example, some tools were designed to both detect and categorise CSAM (Dalins et al. 2018), to detect CSAM by classifying media content and separately the text of file names (Peersman et al. 2016), or to consider both text and images simultaneously (Granizo et al. 2020).

The introduction of AI technologies in place of human decision-making offers important opportunities (Singh & Nambiar 2024). Benefits include automatically classifying CSAM images and improving the efficiency of detection or classification of images and videos when large volumes of data are obtained. These functions have the potential to reduce the risk of psychological harm to investigators. The automated nature of these tools is particularly important given the demands placed on law enforcement by the recent dramatic growth in CSAM production and sharing. The rate of online sharing and viewing of CSAM is currently beyond manual human detection and intervention. Ultimately, this may mean that a human response alone is not an adequate solution to this increasingly AI driven problem, and opportunities to integrate AI technology into existing CSA prevention strategies should be explored.

## Directions for research

The principal gap in research identified by this review was the use of AI for CSA offending. Further, the studies examined indicated that several of the AI technologies proposed for the prevention or disruption of CSA were not fit for purpose in their current form or did not have sufficient evidence to support their efficacy in a real-world setting. Figure 4 provides a summary of important areas for future research at the intersection of AI and CSA, informed by the evidence and gap map in Figure 3.

**Figure 4: Summary of future research required on artificial intelligence technologies in the context of child sexual abuse**

| | |
|---|---|
| | Research into the ways that AI is used to commit CSA offences and evade detection, including but not limited to the aggregation and generation of CSAM |
| | Improving and evaluating existing AI technologies to ensure they are fit for purpose for those who will implement them |
| | Develop nuanced and appropriate ways to test and implement AI technologies in real-world settings—including the development of ethically sourced datasets that could allow for the training and testing of AI tools |
| | Continue to build knowledge on CSA and child sexual offender behaviours and characteristics, to inform effective targeting of newly developed AI technologies |
| | Continued innovation of AI technologies and translation to use for the prevention and disruption of CSA |
| | Interdisciplinary collaboration in the development of AI technologies with a focus on applied use in regulatory and law enforcement settings |

While there is evidence that AI is being used in the process of child sexual offending, particularly CSAM offending, a trend that appears to be expanding (Internet Watch Foundation 2023), there were no identified studies investigating the nature and scope of AI use among offenders. Anecdotal evidence suggests AI is being used to generate deepfake CSAM or could automate cyber-grooming (Butler 2023; Internet Watch Foundation 2023). However, there is limited understanding in the research literature of the scope of this problem or emerging uses of AI for malicious purposes. It is important to implement methods to deter the use of AI for illicit activities. Research should continue to explore the uses of AI among child sexual offenders to better understand the risks posed and ways to address these risks.

While there was a sizable amount of research on the development of AI technologies for CSA prevention and disruption, the research evaluating the efficacy of these models, or how they performed in real-world settings, was limited. It is important to note that tools should only be implemented after their performance has been evaluated. This evaluation should measure the tool's effectiveness (eg accuracy, reliability, specificity), as well as its application (ie whether potential users, such as law enforcement, can employ it and find it helpful). It may also be helpful to measure performance with different data sources or for different tasks, to clarify where models perform well and where they do not.

However, the adoption of AI technology alone for CSA prevention and disruption is unlikely to be a comprehensive solution for either CSA in general or AI produced CSAM. Tech-solutionism, in which technology is implemented as a standalone method of solving a given issue, is often criticised for oversimplifying complex problems, failing to address root causes, and leading to unforeseen negative consequences. Additionally, there are reasonable concerns regarding the use of AI prevention methods without supervision or validation by humans. If AI technology were to be found suitable for use, it should not be adopted at the expense of broader approaches to tackling CSA. Rather, it should be implemented as just one tactic among many to reduce the volume and impact of CSA. Detection and prevention approaches, when featuring AI, should continue to be transparent, should feature a degree of human supervision and should be considered part of a suite of complementary approaches to address CSA, rather than a standalone solution.

Developing and testing AI technologies in the context of CSA prevention may require the use of authentic CSA data sources that reflect real-world settings—for example, CSAM, offender chat logs and police case reports. Of course, there are important ethical and practical implications when building and accessing such data sources. It is appropriate that access to these data are tightly controlled; however, the difficulty accessing data is a significant barrier to assessing whether these AI tools may be useful and implementable. Progress in the field of AI for CSA prevention may therefore require consideration of how researchers can reliably evaluate the efficacy of their tool using appropriate datasets, under agreed upon conditions for access. Additionally, AI models could be advanced by improving knowledge of child sexual offender behaviour (Singh & Nambiar 2024).

AI technology will continue to develop rapidly. While caution should be exercised in implementing these models, innovative approaches to addressing CSA should be encouraged. Each of the directions suggested for future research would substantially benefit from interdisciplinary collaboration, particularly featuring the stakeholders who would ultimately use the technology.

## Conclusion

The research literature has, to date, detailed a range of AI technologies developed with the aim of preventing or disrupting CSA—most commonly, those that detect CSAM or online child sexual offenders. The use of AI to address CSA is an emerging field, and while the evidence for the efficacy of AI technology in this context is limited, the field is moving and developing rapidly. Ultimately, with AI supported CSA offending becoming more widely reported, interest in AI approaches to prevent CSA is likely to grow. This review has emphasised the potential for AI technologies to identify, prevent and disrupt CSA. These technologies offer many advantages but must undergo strict evaluation and safety testing and adhere to ethical protocols before being considered for adoption to complement existing strategies.

## References

*URLs correct as at November 2024*

*Included in review

*Agarwal N, Ünlü T, Wani MA & Bours P 2022. Predatory conversation detection using transfer learning approach. In G Nicosia et al. (eds), *Machine learning, optimization, and data science*. Lecture Notes in Computer Science vol. 13163. Springer International Publishing: 488–499. https://doi.org/10.1007/978-3-030-95467-3_35

*Al-Nabki MW, Fidalgo E, Alegre E & Alaiz-Rodriguez R 2023. Short text classification approach to identify child sexual exploitation material. *Scientific Reports* 13(1): 16108. https://doi.org/10.1038/s41598-023-42902-8

*Anderson P, Zuo Z, Yang L & Qu Y 2019. An intelligent online grooming detection system using AI technologies. *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* New Orleans, LA, USA: IEEE1–6. https://doi.org/10.1109/FUZZ-IEEE.2019.8858973

*Brewer R et al. 2023. Advancing child sexual abuse investigations using biometrics and social network analysis. *Trends & issues in crime and criminal justice* no. 668. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78948

Butler J 2023. AI tools could be used by predators to 'automate child grooming', eSafety commissioner warns. *The Guardian,* 20 May. https://www.theguardian.com/technology/2023/may/20/ai-tools-could-be-used-by-predators-to-automate-child-grooming-esafety-commissioner-warns

*Cardei C & Rebedea T 2017. Detecting sexual predators in chats using behavioral features and imbalanced learning. *Natural Language Engineering* 23(4): 589–616. https://doi.org/10.1017/S1351324916000395

*Cubitt T, Napier S & Brown R 2023. Understanding the offline criminal behavior of individuals who live stream child sexual abuse. *Journal of Interpersonal Violence* 38(9–10): 6624–6649. https://doi.org/10.1177/08862605221137712

*Cubitt T, Napier S & Brown R 2021. Predicting prolific live streaming of child sexual abuse. *Trends & issues in crime and criminal* justice no. 634. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78320

*Dalins J, Tyshetskiy Y, Wilson C, Carman MJ & Boudry D 2018. Laying foundations for effective machine learning in law enforcement. Majura: A labelling schema for child exploitation materials. *Digital Investigation* 26: 40–54. https://doi.org/10.1016/j.diin.2018.05.004

Edwards G, Christensen L, Rayment-McHugh S & Jones C 2021. Cyber strategies used to combat child sexual abuse material. *Trends & issues in crime and criminal justice* no. 636. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78313

Ganann R, Ciliska D & Thomas H 2010. Expediting systematic reviews: Methods and implications of rapid reviews. *Implementation Science* 5: 56–66. https://doi.org/10.1186/1748-5908-5-56

*Gangwar A, González-Castro V, Alegre E & Fidalgo E 2021. AttM-CNN: Attention and metric learning based CNN for pornography, age and child sexual abuse (CSA) detection in images. *Neurocomputing* 445: 81–104. https://doi.org/10.1016/j.neucom.2021.02.056

Garriss K & DeMarco N 2023. FBI warns of using AI deepfakes as part of sextortion schemes. Yahoo! News, 6 July. https://www.yahoo.com/news/fbi-warns-using-ai-deepfakes-212047097.html

*Granizo SL, Valdivieso Caraguay ÁL, Barona López LI & Hernández-Álvarez M 2020. Detection of possible illicit messages using natural language processing and computer vision on Twitter and linked websites. *IEEE Access* 8: 44534–44546. https://doi.org/10.1109/ACCESS.2020.2976530

*Grubl T & Lallie HS 2022. Applying artificial intelligence for age estimation in digital forensic investigations. https://doi.org/10.48550/arXiv.2201.03045

*Guerra E & Westlake BG 2021. Detecting child sexual abuse images: Traits of child sexual exploitation hosting and displaying websites. *Child Abuse & Neglect* 122: 105336. https://doi.org/10.1016/j.chiabu.2021.105336

Henseler H & de Wolf R 2019. Sweetie 2.0 technology: Technical challenges of making the sweetie 2.0 Chatbot. In S van der Hof, I Georgieva, B Schermer & BJ Koops (eds), *Sweetie 2.0: Using artificial intelligence to fight webcam child sex tourism*. Information Technology and Law Series, vol 31. The Hague: TMC Asser Press: 113–134. https://doi.org/10.1007/978-94-6265-288-0_3

High-Level Expert Group on Artificial Intelligence 2019. *A definition of AI: Main capabilities and disciplines*. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

Internet Watch Foundation 2023. How AI is being abused to create child sexual abuse imagery. https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/

*Isaza G, Muñoz F, Castillo L & Buitrago F 2022. Classifying cybergrooming for child online protection using hybrid machine learning model. *Neurocomputing* 484: 250–259. https://doi.org/10.1016/j.neucom.2021.08.148

*Jin P, Kim N, Lee S & Jeong D 2024. Forensic investigation of the dark web on the Tor network: Pathway toward the surface web. *International Journal of Information Security* 23(1): 331–346. https://doi.org/10.1007/s10207-023-00745-4

*Laranjeira C, Macedo J, Avila S & dos Santos JA 2022. Seeing without looking: Analysis pipeline for child sexual abuse datasets. arXiv. Presented at the 5th Conference on Fairness, Accountability and Transparency (FAccT), 2022. https://doi.org/10.48550/arXiv.2204.14110

Long C 2023. First reports of children using AI to bully their peers using sexually explicit generated images, eSafety Commissioner says. ABC News, 16 August. https://www.abc.net.au/news/2023-08-16/esafety-commisioner-warns-ai-safety-must-improve/102733628

Lupariello F, Sussetto L, Di Trani S & Di Vella G 2023. Artificial intelligence and child abuse and neglect: A systematic review. *Children* 10: 1659. https://doi.org/10.3390/children10101659

*Macedo J, Costa F & dos Santos J 2018. *A benchmark methodology for child pornography detection*. Paper presented at the 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images: 455–462. https://doi.org/10.1109/SIBGRAPI.2018.00065

*Meyer M 2015. *Machine learning to detect online grooming* (Master's thesis). https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-260390

Milmo D 2023. AI-created child sexual abuse images 'threaten to overwhelm internet'. *The Guardian*, 25 October. https://www.theguardian.com/technology/2023/oct/25/ai-created-child-sexual-abuse-images-threaten-overwhelm-internet

*Murcia Triviño J, Moreno Rodríguez S, Díaz López DO & Gómez Mármol F 2019. C3-Sex: A chatbot to chase cyber perverts. *2019 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress*: 50–57. https://doi.org/10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00024

Murphy M 2023. Predators exploit AI tools to generate images of child abuse. *Bloomberg*, 23 May. https://www.bloomberg.com/news/articles/2023-05-23/predators-exploit-ai-tools-to-depict-abuse-prompting-warnings#xj4y7vzkg

*Ngejane CH, Eloff JHP, Sefara TJ & Marivate VN 2021. Digital forensics supported by machine learning for the detection of online sexual predatory chats. *Forensic Science International: Digital Investigation* 36: 301109. https://doi.org/10.1016/j.fsidi.2021.301109

*Ngo VM, Gajula R, Thorpe C & Mckeever S 2024. Discovering child sexual abuse material creators' behaviors and preferences on the dark web. *Child Abuse & Neglect* 147: 106558. https://doi.org/10.1016/j.chiabu.2023.106558

Okolie C 2023. Artificial intelligence-altered videos (deepfakes), image-based sexual abuse, and data privacy concerns. *Journal of International Women's Studies* 25(2): 1–16. https://vc.bridgew.edu/jiws/vol25/iss2/11

*Oronowicz-Jaśkowiak W et al. 2024. Using expert-reviewed CSAM to train CNNs and its anthropological analysis. *Journal of Forensic and Legal Medicine* 101: 102619. https://doi.org/10.1016/j.jflm.2023.102619

*Peersman C, Schulze C, Rashid A, Brennan M & Fischer C 2016. iCOP: Live forensics to reveal previously unknown criminal media on P2P networks. *Digital Investigation* 18: 50–64. https://doi.org/10.1016/j.diin.2016.07.002

*Pereira M, Dodhia R, Anderson H & Brown R 2021. Metadata-based detection of child sexual abuse material. https://doi.org/10.48550/arXiv.2010.02387

*Polastro M de C & Eleuterio PM da S 2012. A statistical approach for identifying videos of child pornography at crime scenes. *2012 Seventh International Conference on Availability, Reliability and Security*: 604–612. https://doi.org/10.1109/ARES.2012.71

*Puentes J et al. 2023. *Guarding the guardians: Automated analysis of online child sexual abuse*. https://doi.org/10.48550/arXiv.2308.03880

*Razi A et al. 2023. Sliding into my DMs: Detecting uncomfortable or unsafe sexual risk experiences within Instagram direct messages grounded in the perspective of youth. *Proceedings of the ACM on Human-Computer Interaction* 7: article 89. https://doi.org/10.1145/3579522

*Rodríguez JI, Durán SR, Díaz-López D, Pastor-Galindo J & Mármol FG 2020. C3-Sex: A conversational agent to detect online sex offenders. *Electronics* 9(11): 1779. https://doi.org/10.3390/electronics9111779

*Rondeau J, Deslauriers D, Howard III T & Alvarez M 2022. A deep learning framework for finding illicit images/videos of children. *Machine Vision and Applications* 33(5): 66. https://doi.org/10.1007/s00138-022-01318-6

*Sae-Bae N, Sun X, Sencar HT & Memon ND 2014. Towards automatic detection of child pornography. *2014 IEEE International Conference on Image Processing (ICIP)*: 5332–5336. https://doi.org/10.1109/ICIP.2014.7026079

*Seedall M, MacFarlane K & Holmes V 2019. *SafeChat system with natural language processing and deep neural networks*. https://sure.sunderland.ac.uk/id/eprint/10968/

*Seigfried-Spellar KC et al. 2019. Chat analysis triage tool: Differentiating contact-driven vs. fantasy-driven child sex offenders. *Forensic Science International* 297: e8–e10. https://doi.org/10.1016/j.forsciint.2019.02.028

Singh S & Nambiar V 2024. Role of artificial intelligence in the prevention of online child sexual abuse: A systematic review of literature. *Journal of Applied Security Research* 1–42. https://doi.org/10.1080/19361610.2024.2331885

Thiel D, Stroebel M & Portnoff R 2023. *Generative ML and CSAM: Implications and mitigations*. Internet Observatory Cyber Policy Center, Stanford. https://fsi.stanford.edu/publication/generative-ml-and-csam-implications-and-mitigations

Thorn 2024a. *Introducing Safer Predict: Using the power of AI to detect child sexual abuse and exploitation online.* https://www.thorn.org/blog/introducing-safer-predict-using-the-power-of-ai-to-detect-child-sexual-abuse-and-exploitation-online/

Thorn 2024b. *Youth perspectives on online safety, 2023: An annual report of youth attitudes and experiences.* https://www.thorn.org/research/library/2023-youth-perspectives-on-online-safety/

*Ulges A & Stahl A 2011. Automatic detection of child pornography using color visual words. *2011 IEEE International Conference on Multimedia and Expo*: 1–6. https://doi.org/10.1109/ICME.2011.6011977

*Westlake B et al. 2022. Developing automated methods to detect and match face and voice biometrics in child sexual abuse videos. *Trends & issues in crime and criminal justice* no. 648. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78566

This work was originally published as *Trends & issues in crime and criminal justice* no. 711. It is available online at **https://doi.org/10.52922/ti77802**.

# 8. Drivers and deterrents of child sexual offending: Analysis of offender interactions on the darknet

Heather Wolbers, Timothy Cubitt, Michael John Cahill, Matthew Ball, John Hancock, Sarah Napier and Roderic Broadhurst

**Warning: this chapter contains confronting material.**

Sexual abuse of children is a global problem with devastating impacts. Recently, research into the drivers of offending and opportunities for deterrence among online and contact child sexual offenders (CSOs) has proliferated. For example, viewing child sexual abuse material (CSAM) has been associated with increased risk of approaching children online and perpetrating contact sexual abuse, highlighting the importance of disrupting online activities (Insoll, Ovaska & Vaaranen-Valkonen 2021). Research has found that offenders network online and teach one another how to abuse children (Huikuri 2022). To inform approaches aimed at disrupting the occurrence and recurrence of child sexual abuse, the current study examined the drivers and deterrents of offending among a group of undetected online and contact CSOs on the darknet.

While many online CSOs do not commit contact sexual offences against children, and some contact offenders never offend online, available literature suggests a level of overlap between online and contact offending (eg Cubitt, Napier & Brown 2022; Teunissen & Napier 2023). For example, several reviews have found that a small proportion of online sexual offenders will subsequently be detected committing a contact sexual offence (Dowling et al. 2021; Seto & Eke 2015), although this proportion may be higher among offenders who have not yet been detected. Importantly, emerging evidence suggests that viewing CSAM may be an antecedent for contact sexual abuse in some individuals (Insoll, Ovaska & Vaaranen-Valkonen 2021).

CSOs are responsible for devastating outcomes for victims, including mental illness, substance abuse, revictimisation and offending in adulthood (Cashmore & Shackel 2013; Ogloff et al. 2012). Victims whose abuse is recorded and shared online also experience helplessness, shame, anxiety, grief, betrayal and revictimisation (Canadian Centre for Child Protection 2017). Preventing the sexual abuse of children is vital, and a key step in prevention is detecting CSOs and applying suitable and effective inventions. However, offenders have increasingly used new technologies, such as the darknet, to avoid detection (Blokland et al. 2024).

## Child sexual offending on the darknet

The darknet not only impedes the detection of child sexual offending; it actively enables offending. The darknet is a private network where connections are made using non-standard protocols and ports, creating a space where users can operate with a high level of anonymity and privacy, attracting cybercriminal activities including child sexual offending.

In 2015, there were 900 child sexual abuse forums active on the darknet (Owen & Savage 2015). While these forums made up just two percent of all active darknet sites, they accounted for 80 percent of the daily browsing requests. The darknet enables viewing and trading of CSAM with limited risk of discovery (Huikuri 2022). Further, some online forums boast large international memberships and an overwhelming amount of information, media and contacts (Connolly 2021). These forums offer a sense of community and support, promote pro-offending beliefs and perpetuate justifications for offending (Cantor et al. 2022; O'Halloran & Quayle 2010). Moreover, forums can facilitate sexual violence and offer advice on how to offend and conceal offending against children (Huikuri 2022; van der Bruggen & Blokland 2021).

### Studying undetected child sexual offenders

There are several barriers to detecting CSOs, the first of which is reporting. Rates of reporting of sexual violence are low, and under-reporting is more pronounced when the victim is a child (Scurich 2020). Investigation, arrest and prosecution rates for sexual offending are even lower, meaning a large proportion of CSOs go unreported and unidentified. This has implications for how we understand CSOs—if we can only measure the characteristics and behaviours of those who have been detected by authorities, our knowledge will be incomplete as we cannot be sure if those who remain undetected are similar or entirely different.

It is pivotal to better understand CSOs who go undetected and to develop strategies to improve the visibility of this group and subsequent detection and intervention opportunities. We therefore ask one principal research question: among undetected CSOs on the darknet, what are the drivers and deterrents of sexual offending against children?

## Method

### Data source

Data for this research were scraped from a darknet forum for individuals who self-identified as a paedophile and held 'conservative' views of sexual contact with children, which is a key point of difference from other studies examining CSO darknet forums (eg van der Bruggen & Blokland 2021). One forum user described being conservative to mean:

> … someone who does not have intimate sexual relations, or engage in "inappropriate touching", with children because of moral and/or ethical reasons (User 88, Thread 3)

In other words, this forum featured individuals who claimed not to engage in contact offences against children but who viewed non-contact child sexual offences as comparatively acceptable. Despite the expressed conservative focus, some forum users admitted to prior contact offences and some believed they would engage in future contact offences. Further, users frequently discussed engaging with CSAM. In this chapter, we refer to forum users collectively as CSOs. While this term usually refers to those who have been convicted of a child sexual offence, it is appropriate for this group due to the nature of the forum and the self-reported offending histories of the users.

The data source is unique as it offers insight into CSOs who have largely remained undetected by law enforcement, family or friends. It also allows insight into the thoughts of CSOs when in conversation with one another, in a setting where they feel reasonably comfortable. Users had confidence in their anonymity because the forum was on the darknet. This anonymity and the community that developed on the forum appeared to encourage open discussions between self-identifying CSOs.

## Sample

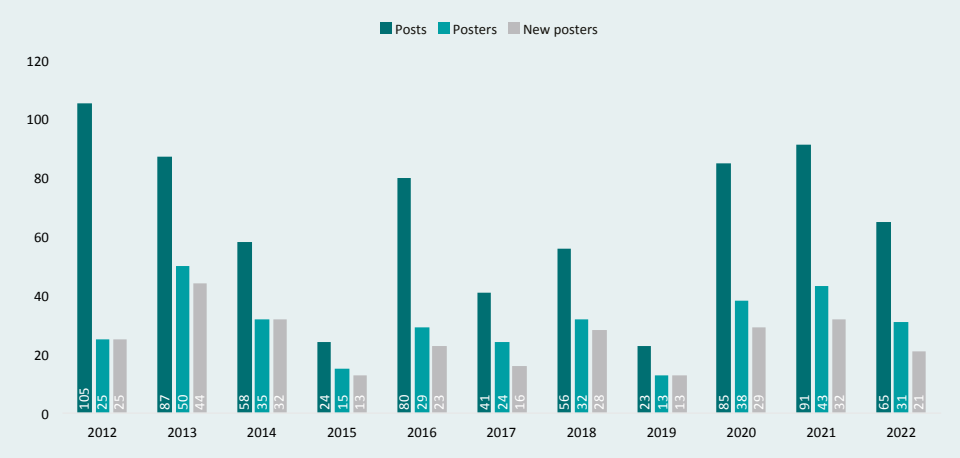
This research focused on posts relating to help-seeking or accessing support, with 17 forum threads identified as thematically relevant (Table 1). To adhere to the conditions of the Human Research Ethics Committee's approval, members of the research team who scraped the forum received guidance from government and law enforcement in relation to the handling of the content. These members identified and removed any content deemed written CSAM, avoiding exposure of the remaining research team. To gain access to the forum, the research team made several posts including an introduction. These were also removed prior to analysis.

**Table 1: Summary of darknet forum threads**

| No. | Paraphrased thread topic | Posts |
|---|---|---|
| 1 | Would you take a magic pill to lose sexual attraction to children? | 289 |
| 2 | How do you stop yourself? | 69 |
| 3 | I want to quit my use of CP [child pornography] and I hope you can offer me some support. | 48 |
| 4 | What do you think about the DDlg [daddy dom/little girl] community? | 44 |
| 5 | I would like to share what I have learned from therapy. | 33 |
| 6 | What do you think about 3D animation? | 33 |
| 7 | I have switched to fiction to break my CP habit. Has anyone else tried this? | 31 |
| 8 | Sexual contact with a child may or may not be damaging. Avoid contact for the child's sake. | 28 |
| 9 | I am struggling with temptation. Does anyone have ideas to help? | 22 |
| 10 | Would you take a pill to remove your libido? | 21 |
| 11 | What support is available? | 20 |
| 12 | I feel beyond redemption or forgiveness. | 18 |
| 13 | Conservative pedophile vs not, selfish vs unselfish, evil vs not. | 17 |
| 14 | An obituary for my collection | 16 |
| 15 | What I want clashes with my morals and stops me from developing a strong attraction to adults. Has anyone been in a similar situation and can share their story? | 14 |
| 16 | I'm getting help. What strategies have you learned from professionals? | 7 |
| 17 | I'm terrified to open up to my therapist. What would happen? | 5 |

In total, 715 posts from February 2012 to June 2022 were examined (Figure 1). The annual number of posts on the relevant threads fluctuated over time, with peaks in 2012, 2016, and 2020 to 2022. The latest peak was likely explained by the COVID-19 pandemic, in which there was a suspected increase in online child exploitation (Salter & Wong 2021). While the number of unique forum users between February and July 2012 could not be determined due to archiving (76 posts), there were 276 unique users active between August 2012 and June 2022 (639 posts), most of whom had not previously posted on the threads we examined. These 276 unique users were each assigned a number, while forum users posting before August 2012 are referred to as 'unknown users'. On average, users made 2.3 posts on the examined threads (range: 1–34). Importantly, we could not estimate the number of individuals reading the forum without posting, or the number of individuals on the forum beyond the 17 threads examined; however, estimates of this type could be useful for future research.

**Figure 1: Number of forum posts and users per year**




Note: The number of users and new users in 2012 does not include those prior to August. 'New posters' refers to posters who had not previously posted on the threads we examined.

Users typically did not reveal identifying information about themselves, meaning there was little information on their demographics. However, it was clear that those who posted on the forum were primarily male, with a few females who identified themselves as being in a sexual relationship with, or married to, a CSO. The location of forum users was rarely mentioned, but there were indications that users were from a range of countries and, for several, English was not their first language.

## Analytical approach

The forum posts were coded inductively rather than deductively using a grounded theory approach (Strauss & Corbin 1998). This approach focuses on the views of research participants, allowing themes to be developed from posters' thoughts and opinions. Grounded theory analysis involves three steps:

- open coding—reading the raw data and considering preliminary codes and themes;
- axial coding—identifying relationships, similarities and differences in the preliminary codes and themes; and
- selective coding—reviewing the raw data, codes and themes to identify and consider overarching variables that describe connections and contrasts within the data.

A coding framework was developed independently by two members of the research team, with two segments of data double coded to assess interrater reliability. Early in the coding process, a match rate of 61.3 percent was achieved, and differences in coding were discussed and addressed. The second segment then had an improved match rate of 84.7 percent.

## Limitations

The data were sourced from a single forum on the darknet, meaning findings may not be representative of the wider content or user groups found on other darknet forums. Further, because we examined a subset of threads from the forum, and certain posts were excluded prior to analysis, it is possible that this affected how representative the remaining posts are of the wider forum. As a result of these limitations, we are cautious not to generalise the findings beyond the selected threads and sample of individuals who participated in these threads.

While forum users self-identified as paedophiles, this was a self-imposed label. With the information available, we were not able to verify whether objective diagnostic criteria for paedophilia were met. We can be certain, though, that when forum users referred to children they meant individuals under the age of 18, and we can be confident from the content of their conversations and their use of the darknet that they perceived their behaviour to be illegal. However, the exact age range to which they referred may have differed between users, and this may have influenced the context of their conversations.

# Results

There were a range of overarching themes identified in the data, summarised in Table 2. These themes could broadly be described as drivers and deterrents of child sexual offending. Each of these can be broken down into two categories: those relating to how forum users perceived themselves and their offending, which we refer to as cognitive drivers or deterrents, and social or environmental factors, which we refer to as socio-environmental drivers or deterrents.

| Table 2: Drivers and deterrents of child sexual offending emerging from offender interactions on the darknet | |
|---|---|
| **Drivers of offending** | |
| Cognitive drivers | Normalising sexual offences against children |
| | Shifting responsibility and denying the impact of offending |
| | Minimising the harm of child sexual abuse material |
| | Denying the capacity for self-control |
| Socio-environmental drivers | Online communities |
| | Technology |
| | Adult partners in a sexual relationship with an offender |
| **Deterrents of offending** | |
| Cognitive deterrents | Harm to children and the shame attached to child sexual offences |
| Socio-environmental deterrents | Detection by friends, family or authorities |

## Cognitive drivers of offending

Four cognitive drivers of child sexual offending were identified among the darknet forum posts. These influenced how CSOs viewed their offences and the resulting impacts on children, and alleviated their perceived responsibility for their behaviour.

### *Normalising sexual offences against children*

For some, the pathway towards identifying as a CSO featured a prolonged process of accepting and normalising sexual offences against children. The culture of the forum supported this normalisation, with some users encouraging others to consider it a point of pride. One user describes being a CSO as 'A gift, rather than a curse' (Unknown user, Thread 1).

> you must be confident in your self, you need to accept who you are and make this a good thing, if you are a pedophile - it's great! It's your unique and wonderfull path that you need to carefully but surely follow. (User 270, Thread 15)

Some CSOs acknowledged conventional morality, while simultaneously prioritising their identity as a CSO over the potential impact on children.

> I spent two decades being anti-pedo, thinking I was a monster, then practically in an instant, became pro-pedo. Life is better now I have stopped hating my own nature. My epiphany wasn't that my moral code changed. It was an intellectual breakthrough: a realization that I had been mis-applying my moral code (User 264, Thread 15)

### *Shifting responsibility and denying the impact of offending*

A common belief shared by forum users was that sexual contact with children was not harmful, that children could consent to and be an equal partner in a sexual interaction with an adult. These discussions elevated the role of the child to that of an offender's equal and, in doing so, minimised the harm to the victim.

> I feel as though I can understand the sexual nature of children better than a non-pedophile would. I'm not neglectful of the fact that sexual touches to children can be beneficial and healthy as opposed to detrimental. (User 191, Thread 1)

Many opposed what they considered to be the mainstream view that children were victims and believed their view of children was superior. Rather than being ignorant to the harm of sexual offences against children, they outright denied that sexual offences were the cause of the harm. This view shifted the responsibility for harming children from the offender to society.

> Is a child's innocence stolen when they are having mutually enjoyable sexual contact with an adult, or was it stolen when outside parental and/or authority figures tell them it was wrong and evil? (Unknown user, Thread 8)

Shifting responsibility, and reframing the morality of sexual offences against children, was a key cognitive driver among CSOs. Importantly, these views were reinforced by others, creating an environment that alleviated responsibility from offenders, allowing them to rationalise or justify their behaviour.

### *Minimising the harm of child sexual abuse material*

While forum users made efforts to reframe responsibility for and impact of contact offending, the same was true for CSAM. In particular, those who discussed CSAM actively minimised the harm done to children.

> I'm just looking at bunch of pixels on my computer. I'm not hurting anybody. (User 39, Thread 2)

Forum users often made distinctions between viewing CSAM and contact offending with the intention of minimising the impacts of viewing CSAM and normalising it as an alternative.

> we all know that CP [child pornography] reduces sex offenses with actual children (User 8, Thread 6)

Some took this notion a step further, positioning themselves as morally superior among CSOs because they viewed CSAM instead of contact offending.

> I somehow justify as long as I only look it is no where near as bad as the person that I have seen pics or vids of touching or more. (User 41, Thread 2)

Forum users minimised the harm of CSAM and its connection to contact offending. This minimisation of harm and the positioning of some child sexual offenders as morally superior to others appeared to be strong cognitive drivers of offending.

### *Denying the capacity for self-control*

While a large proportion of forum users were unwilling to take ownership of the harm associated with their offending, those who did often described what they viewed as an inability to implement self-control. These forum users felt their lack of self-control meant they were not responsible for their behaviour and should not be held accountable for the resulting harm. This acted as a self-serving mechanism supporting their offending.

> What I was confused about was why I was sexually attracted to little girls … For me that desire was overwhelming and uncontrollable. It caused me to make some dangerous decisions and mistakes. (User 267, Thread 15)

At times, a lack of self-control was attributed to hyper-sexual arousal. In settings featuring sexual arousal, they distanced themselves from responsibility for their behaviours.

> I wouldn't say that I am especially aroused by "hurtcore fantasies", it is just that in times of arousal I lose most of my moral boundaries. It is not that that kind of stuff gets me going, but I don't mind it at that time until I come back to my senses. (User 46, Thread 9)

When discussing CSAM offending, forum users again denied responsibility, likening their behaviour to an addiction.

> it's an addiction that IMO [in my opinion] is as powerful as any drug, and without the side effects of any crash. (User 180, Thread 3)

Cognitive drivers of offending expressed on this forum demonstrated the complex internal justification for offending among CSOs. In particular, forum users attempted to normalise their preferred offending and diminish the harm it caused, shift responsibility and deny their capacity for self-control. However, there were important and connected socio-environmental drivers that also supported offending.

## Socio-environmental drivers of offending

### *Online communities*

Users often referred to the forum as a community of individuals who supported each other, suggesting that, while users self-identified as CSOs, they also identified as a community rather than a disparate group of like-minded individuals.

> I'm just glad that you have a place like this to gain advice and guidance at your age. Most of us didn't have that option. I for one am grateful for this community. (User 57, Thread 9)

Online communities and networks appeared to help shape views about and alleviate responsibility for offending. Forum users explained the importance of engaging with this online community in reshaping their view of child sexual offending and, pivotally, increasing their likelihood of continued offending.

> While I FEEL committed to this and have tools I've developed to stay so, this place is making me reconsider, insomuch as I'm not really sure how damaging a sexual relationship would be to a 15yo girl. (User 56, Thread 9)

While many forum users engaged in discussions designed to passively alleviate their responsibility for harming children, others directly and explicitly encouraged sexual offences against children.

> If you're not harming anyone and yourself, why not? There's no overlord holding rule over your decisions. (User 54, Thread 9)

### *Technology*

At times forum users mentioned using technologies to enable and conceal their offending. Tor networks, secure storage and erasure programs offered increased opportunities to offend and avoid detection, driving continued offending.

> I know that obsession, went through it when I first discovered the darknet … even with TOR, I'm not convinced that seeking out material and then deleting it after use is actually more secure than keeping a small store, well encrypted (in a "plausible deniability" way). (User 54, Thread 14)

Users sought guidance about reliable and covert methods of collecting CSAM. Often, this involved a planning period and gathering technological resources prior to offending.

> For many years, I wanted to start [a collection] but did not have the resources to do so. It took great time and planning to see him come to fruition. I sought out the counsel of my peers and took great time and effort to provide him a place to grow and fulfill his potential. (User 24, Thread 4)

Despite the availability of technology to help conceal a CSAM collection, for some the risk was perceived as too great, leading to technology-enabled offending approaches with a smaller online imprint. Many users discussed cycles of downloading or viewing CSAM, then erasing traces of offending. This appeared to be a common pattern used when attempting to avoid detection.

> I adhere to "download, fap, purge" pattern using the eraser program … (User 139, Thread 2)

### *Adult partners in a sexual relationship with an offender*

At times, forum users discussed sexual interactions with other adults that simulated adult–child sexual contact. For example, there was a discussion about a role-play style referred to as 'daddy dom/little girl' (DDLG), in which CSOs engaged in sexual activity with an adult role-playing as a child.

> I absolutely love DDLG … I feel doing it keeps my desire to children at bay, as I get my fix when I'm fucking a woman. I love when it becomes rape play too (User 223, Thread 4)

For some users, adult sexual relationships involving this type of role-play were motivated by the need to avoid detection or consequences. However, others were acutely aware of the perception that this type of role-play was associated with CSOs, and viewed it as holding some risk.

> The main way I go about indulging fantasies is online roleplays … I like the back and forth, and the fact that all of those fantasies are being fulfilled in some way with absolutely no consequence. (User 186, Thread 4)

Separate to those who identified as CSOs were self-identifying female forum users. These users described their experiences of a sexual relationship with a partner who they knew to be a CSO.

> My first DD was a no-contact pedophile, and I took great pride in being able to help him satisfy his desires in a safe and legal way … Both my DDs enjoyed looking at pictures of me from my childhood as a more tangible reference for the little girl who still lives inside my heart. (User 228, Thread 4)

These forum users believed that, by supporting their partner, they were reducing the likelihood of contact offending. They went to considerable lengths to divert their partner from contact offending, with one user reporting she had accessed CSAM with her partner.

> With my first DD, we would sometimes look at CP together. (User 228, Thread 4)

Forum users considered role-play interactions an alternative to sexual contact with a child and therefore a deterrent of offending. Importantly, these interactions normalised and reinforced the sexualisation of children and child sexual offending, and at times involved accessing CSAM, consequently supporting further offending. However, it must be acknowledged that these women in sexual relationships with CSOs were very unusual in the context of available research. Typically, findings relating to partners show that they experience coercive control and at times physical and sexual violence in their relationships with offenders (Jones, Woodlock & Salter 2021; Salter, Woodlock & Dubler 2022). However, these findings suggest that the partners in our sample accessed a darknet forum for CSOs to discuss attempts to divert their partners' offending. We must therefore acknowledge that, while some of their behaviours may have supported offending, these individuals are most likely not representative of the majority of partners.

## Cognitive deterrents of offending

### *Harm to children and the shame attached to child sexual offences*

While many users attempted to dissociate themselves from the harm caused by child sexual offences, others acknowledged the harm of these offences, although this perspective was less common among the threads analysed. In particular, awareness of these harms was sometimes discussed as a deterrent to offending.

> I refuse to be the monster in another adults life long story of abuse. I refuse to be the selfish one who tricked a kid into doing something they didnt understand and couldnt understand. (User 201, Thread 12)

However, there were diverse opinions on the forum. While some users considered the harm experienced by children as unacceptable, others centred their reasons for not offending upon themselves, noting that they experienced guilt and shame.

> However much I got off scot-free this time, the fear is that next time it will end up much worse. Not in terms of getting caught and ending up in jail, but actually hurting someone needlessly. (User 208, Thread 16)

## Socio-environmental deterrents of offending

### *Detection by friends, family or authorities*

A small number of forum users discussed being detected by family or friends. One user described arousing suspicion among friends, who began questioning their involvement with children. This was an example of detection, even when it is not by authorities, strongly deterring offending behaviour.

> I've played my part of pushing the button. At first it went well, but then my friends saw through me. Since I first brought up the idea that pedophiles are just as much humans as the next person, my motivation to be around children has been harshly critisized. (Unknown user, Thread 1)

Another user was detected by their partner, resulting in a confrontation. While this user discussed the shame attached to being detected by their partner, they did not change their behaviour, explaining that detection by family was not sufficient to change their behaviour.

> When I was unable to accept myself as the person I am, I came out to my wife. (That magnificent woman forced me into a corner. I had no choice but to be honest with her. There was no courage in my admission to her.) She began in me the process of forcing me to be honest with myself. (Unknown user, Thread 1)

For many, concern about being detected by the people in their lives was enough to influence offending behaviour.

> The worst thing I could imagine is that I am found out. Nearly all of my friends abandon me. My wife and kids—gone. Forever. My life would be over as I know it … It really pains me when I think of where I cannot go, what I keep inside, and how I must behave around nearly every other person I contact. (User 10, Thread 1)

While detection by family or friends appeared to have a varying impact on users, CSOs overwhelmingly expressed a fear of being caught by law enforcement. This fear played a strong role in behaviour regulation and the likelihood of offending. For some, the fear of being caught either resulted in either total desistance or prevented certain types of offending deemed too risky.

> Even though I have had opportunities with little girls … I chose not to follow through that could ultimately get me in trouble with the law that will ruin my life. (User 139, Thread 2)

# Discussion

## Drivers of child sexual offending

Consistent with prior research, this study found many of the drivers of child sexual offending involved biased thinking which at times supported ongoing offending (Cantor et al. 2022; Insoll, Ovaska & Vaaranen-Valkonen 2021; O'Halloran & Quayle 2010). Forum users developed distorted perceptions of the world, victims or themselves, designed to normalise or deny the impact of CSAM and sexual contact with children. Simultaneously, they shifted the responsibility for their offences away from themselves by denying their capacity for self-control and blaming society for the harm done to victims. Forum users supported, workshopped or adopted the perspectives of others, which suggests the darknet forum environment contributed to the development of cognitive drivers for offending. Forum interactions appeared to lead users to find a perspective they considered morally acceptable, limiting perceived responsibility for their own behaviour and consequently driving offending. Child sexual offences were also supported and reinforced by other online and offline communities. For example, forum users described adult sexual relationships involving adult–child role-play, which desensitised CSOs to the sexualisation of children and normalised child sexual abuse.

Prior research suggests contact with other offenders is associated with seriousness of offending, such as escalation from collecting to trading and distributing CSAM (Carr 2007; Merdian et al. 2013). Our findings suggest that the ability to share information on offending and avoiding detection was a key driver of offending among forum users. For example, they discussed ways of using technology to facilitate and conceal offending by enabling the access to, storage and erasure of CSAM while maintaining privacy and anonymity. Pivotally, research suggesting CSOs use few technological measures to avoid detection has typically studied offenders who had been apprehended (Balfe et al. 2015). Our findings, focusing on those who had not yet been detected, highlight the role of technology in avoiding detection. While this finding reinforces the importance of studying undetected CSOs, it also suggests past research may have underestimated the sophistication of technology use among offenders.

## Deterrents of child sexual offending

Fear of being caught by law enforcement was prominent among forum users. While this alone did not completely deter offending, forum users reported changing their behaviour to avoid activity that posed a high risk of attracting law enforcement attention. Similarly, detection by family and friends was a notable concern, motivating CSOs to moderate their behaviour. However, among those who reported being detected by family or friends, the discovery did not appear to trigger sustained desistance.

Finally, while many forum users held distorted views that sexual abuse does not harm children, these views were not universal. Some acknowledged the harm associated with offending and expressed deep shame and guilt. This appeared to meaningfully deter contact offending and, to a lesser extent, CSAM use.

## Implications

Our findings suggest that when CSOs can freely communicate and share information with one another, they may reinforce distorted perceptions of their behaviour and impacts on victims, thereby driving offending. These findings align with those of prior research suggesting that CSOs prefer to justify their behaviour, and deny its harm, rather than excusing it (Durkin & Bryant 1999). Importantly, on forums such as these, there is little resistance to ideas that support offending. Even when forum users raised potential risks (such as detection by law enforcement), the discussion included methods of reducing risk while continuing to offend. These findings support approaches that aim to disrupt information sharing between CSOs and their use of technology. Such approaches could limit the drivers of offending among undetected offenders (Leclerc et al. 2021).

Notably, this research identified several independent but related cognitive drivers of offending. While darknet research into CSOs is emerging, these findings suggest that psychosocial intervention approaches may be useful in helping to limit the development and normalisation of beliefs that support offending. However, the deterrents of offending also offer opportunities to reduce risk among CSOs. For example, fear of detection by law enforcement was a strong regulator of behaviour, highlighting the importance of continued disruption efforts by law enforcement. Forum users demonstrated clear concern about law enforcement detection, resulting in their offending becoming increasingly tech enabled, and interactions moving to the darknet. Evidently, law enforcement has a key role to play in detecting and apprehending CSOs, and collaborating with tech companies and internet service providers to increase the visibility of offenders and limit their ability to interact on the darknet.

## References

*URLs correct as at July 2024*

Balfe M et al. 2015. Internet child sex offenders' concerns about online security and their use of identity protection technologies: A review. *Child Abuse Review* 24(6): 427–439. https://doi.org/10.1002/car.2308

Blokland A et al. 2024. Why do users continue to contribute to darknet child sexual abuse material forums? Examining social exchange, social capital, and social learning explanations using digital forensic artefacts. *Child Abuse & Neglect* 154: 106815. https://doi.org/10.1016/j.chiabu.2024.106815

Canadian Centre for Child Protection 2017. *Survivors' Survey: Full report 2017*. https://protectchildren.ca/en/resources-research/survivors-survey-results/

Cantor N et al. 2022. A content analysis of posts to an online support forum for "girl lovers". *Journal of Child Sexual Abuse* 31(5): 522–537. https://doi.org/10.1080/10538712.2022.2112348

Carr A 2007. *Internet censorship offending: A preliminary analysis of the social and behavioural patterns of offences* (Doctoral thesis). Bond University, Gold Coast, Australia. https://research.bond.edu.au/en/studentTheses/internet-censorship-offending-a-preliminary-analysis-of-the-socia

Cashmore J & Shackel R 2013. *The long-term effects of child sexual abuse*. Child Family Community Australia paper no. 11. https://aifs.gov.au/resources/policy-and-practice-papers/long-term-effects-child-sexual-abuse

Connolly K 2021. German police shut one of world's biggest darknet child abuse images sites. *The Guardian*, 3 May. https://www.theguardian.com/world/2021/may/03/german-police-shut-one-of-worlds-biggest-darknet-child-abuse-images-sites

Cubitt TIC, Napier S, Brown R 2022. Understanding the offline criminal behavior of individuals who live stream child sexual abuse. *Journal of Interpersonal Violence* 38(9): 6624–6649. https://doi.org/10.1177/08862605221137712

Dowling C, Boxall H, Pooley K, Long C & Franks C 2021. Patterns and predictors of reoffending among child sexual offenders: A rapid evidence assessment. *Trends & issues in crime and criminal justice* no. 632. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78306

Durkin KF & Bryant CD 1999. Propagandizing pederasty: A thematic analysis of the on-line exculpatory accounts of unrepentant pedophiles. *Deviant Behaviour* 20(2): 103–127. https://doi.org/10.1080/016396299266524

Huikuri S 2022. *The darkest side of the darknet: How do online communities of pedophiles contribute to the justification of sexual violence against children?* Reviews of the Police University College 25. Tampere: Police University College. https://urn.fi/URN:NBN:fi-fe2022082656518

Insoll T, Ovaska A & Vaaranen-Valkonen N 2021. *CSAM users in the dark web: Protecting children through prevention*. Suojellaan Lapsia (Protect Children). https://www.suojellaanlapsia.fi/en/post/csam-users-in-the-dark-web-protecting-children-through-prevention

Jones C, Woodlock D & Salter M 2021. *Evaluation of PartnerSPEAK*. Sydney: UNSW. https://www.academia.edu/60971032/Evaluation_of_PartnerSPEAK

Leclerc B et al. 2021. Child sexual abuse material on the darknet: A script analysis of how offenders operate. *Trends & issues in crime and criminal justice* no. 627. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78160

Merdian HL, Curtis C, Thakker J, Wilson N & Boer DP 2013. The three dimensions of online child pornography offending. *Journal of Sexual Aggression* 19(1): 121–132. https://doi.org/10.1080/13552600.2011.611898

O'Halloran E & Quayle E 2010. A content analysis of a "boy love" support forum: Revisiting Durkin and Bryant. *Journal of Sexual Aggression* 16(1): 71–85. https://doi.org/10.1080/13552600903395319

Ogloff J, Cutajar M, Mann E & Mullen P 2012. Child sexual abuse and subsequent offending and victimisation: A 45 year follow-up study. *Trends & issues in crime and criminal justice* no. 440. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti251190

Owen G & Savage N 2015. *The tor dark net*. Global Commission on Internet Governance paper no. 20. https://researchportal.port.ac.uk/en/publications/the-tor-dark-net

Salter M & Wong WKT 2021. *The impact of COVID-19 on the risk of online child sexual exploitation and the implications for child protection and policing*. University of New South Wales. https://www.end-violence.org/knowledge/impact-covid-19-risk-online-child-sexual-exploitation-and-implications-child-protection

Salter M, Woodlock D & Dubler N 2022. Secrecy, control and violence in women's intimate relationships with child sexual abuse material offenders. *Trends & issues in crime and criminal justice* no. 661. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78856

Scurich N 2020. Introduction to this special issue: Underreporting of sexual abuse. *Behavioral Sciences & the Law* 38(6): 537–656. https://doi.org/10.1002/bsl.2491

Seto MC & Eke AW 2015. Predicting recidivism among adult male child pornography offenders: Development of the Child Pornography Offender Risk Tool (CPORT). *Law and Human Behaviour* 39(4): 416–429. https://doi.org/10.1037/lhb0000128

Strauss A & Corbin J 1998. *Basics of qualitative research: Techniques and procedures for developing grounded theory*, 2nd ed. Sage Publications

Teunissen C & Napier S 2023. The overlap between child sexual abuse live streaming, contact abuse and other forms of child exploitation. *Trends & issues in crime and criminal justice* no. 671. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78993

van der Bruggen M & Blokland A 2021. A crime script analysis of child sexual exploitation material fora on the darkweb. *Sexual Abuse* 33(8): 950–974. https://doi.org/10.1177/1079063220981063

This work was originally published as *Trends & issues in crime and criminal justice* no. 703. It is available online at **https://doi.org/10.52922/ti77659**.

# 9. Financial risk indicators of child sexual abuse live streaming: A proof of concept prediction model

Timothy Cubitt, Sarah Napier and Rick Brown

Live streaming of child sexual abuse (CSA) is a technologically facilitated crime characterised by the procurement and viewing of sexual abuse of children across the internet, in real time, in exchange for money (Açar 2017; Europol 2019). This crime type is often arranged and facilitated by a third party (Napier, Teunissen & Boxall 2021; Ramiro et al. 2019; Terre des Hommes 2013) and can feature an offender directing the abuse via the internet. The tech-enabled and online nature of CSA live streaming results in considerable barriers to monitoring by authorities, with little tangible evidence of the offence beyond a financial transaction and metadata (Açar 2017). Some evidence on the characteristics of financial transactions by CSA live streaming offenders has emerged (AUSTRAC 2022; Brown, Napier & Smith 2020; Cubitt, Napier & Brown 2023, 2021; Financial Action Task Force 2025; Gibraltar Financial Intelligence Unit nd). However, we do not yet know whether these transactions are meaningfully different from other financial transactions.

## What is known about live streaming of child sexual abuse

CSA live streaming is also known as 'webcam child sex tourism/abuse' (Masri 2015; Puffer et al. 2014; Terre des Hommes 2013) and 'live-distance child abuse' (AUSTRAC 2019; EFC 2015). The Philippines has been identified as an epicentre of CSA live streaming offences (AUSTRAC 2019; ECPAT International 2017; Europol 2019; Puffer et al. 2014), facilitated by high-speed internet and good English language proficiency (Batha 2016; ECPAT International 2017; Puffer et al. 2014). A key factor in CSA live streaming offences is the interaction between those who procure the offences and those who facilitate them (Açar 2017; ECPAT International 2017; Europol 2019; Global Alliance against Child Sexual Abuse Online 2016). Facilitators of CSA live streaming are the recipients of financial transactions procuring the abuse, who then arrange the live streaming session on behalf of the offender. However, it is important to note that, in some instances, offenders also procure CSA live streams directly from victims (Napier, Teunissen & Boxall 2021).

There is little evidence that the live streaming of CSA is decreasing over time. Europol (2020a) reported an increase in sharing of child sexual abuse material (CSAM) on the darknet, including content classed as 'live streams' captured via webcam. A recent estimate suggested that, in 2022 alone, nearly one in every 100 Filipino children had experienced trafficking to produce CSAM, including CSA live streaming, and nearly three in every 1,000 adults had been involved in producing CSAM for profit (International Justice Mission 2023).

### The financial element of child sexual abuse live streaming

Unlike some other forms of online sexual exploitation of children (DeHart et al. 2017; Napier et al. 2024), CSA live streaming typically involves a financial transaction. The technological and financial elements of CSA live streaming result in both challenges and opportunities for disruption by law enforcement. On the technological side, the evidence left by CSA live streaming offences is difficult to obtain—typically only resulting in streaming logs, chat logs and metadata (Açar 2017; ECPAT International 2017; Europol 2020b; Netclean 2019)—meaning the live streaming event is difficult to detect with this evidence alone. However, there appear to be additional detection opportunities on the financial side of live streaming offences. Police often rely on money transfers, call histories and chat logs for evidence during investigations (ECPAT International 2017), while early evidence has suggested that the characteristics of money transfers for procuring live streaming may be consistent enough to detect offending (Brown, Napier & Smith 2020; Cubitt, Napier & Brown 2023, 2021).

The present research uses a case-control design to consider the differences in the characteristics of financial transactions between those who make payments for live streaming of CSA and those who do not. We then analyse the financial characteristics of CSA live streaming to produce financial risk indicators. This research is oriented toward practice, so we undertake this analysis to understand whether these types of approaches could be useful in supporting disruption and intervention with perpetrators.

## Method

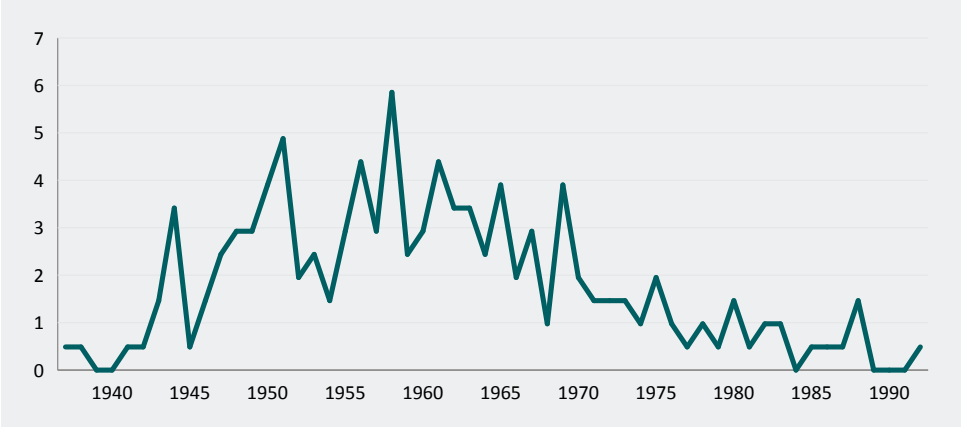
### Data

To detect and prevent financial crime, the Australian Transaction Reports and Analysis Centre (AUSTRAC) collects financial transaction data relating to individuals in Australia. Separately, the Australian Criminal Intelligence Commission (ACIC) collects criminal history information relating to individuals who have come to police attention for offending in Australia. In 2018, the Philippine National Police and the Philippine National Bureau of Investigation provided the Australian Federal Police with a list of 118 individuals arrested for facilitating CSA. AUSTRAC used this information to identify 299 Australia-based individuals who had made financial transactions with these known CSA facilitators. AUSTRAC provided the ACIC with information on the 2,714 transactions made by these Australia-based individuals, and the ACIC then linked the data with criminal history information using names and dates of birth, resulting in a deidentified dataset of 256 individuals. (See Brown, Napier & Smith 2020 for the full data-matching methodology.)

After the data were gathered and matched, the Australian Institute of Criminology took receipt of the data for analysis. These data consisted of transactions processed between January 2006 and February 2019, limited demographic information, and criminal histories. Consistent with prior research (Cubitt, Napier & Brown 2023, 2021), there were cases in which demographic data were not available and these were excluded from analysis prior to undertaking matching. AUSTRAC then provided a matched dataset of individuals who had not been identified for making a transaction for CSA live streaming, for comparison.

There were three pieces of information available to use in matching the comparison data: year of birth (see Figure 1), whether they had transferred money from Australia to the Philippines, and whether they had done so within the date range of the individuals who transacted with facilitators (January 2006 to February 2019). In instances where the data did not contain demographic details of members of the live streaming group, they were excluded from analysis as we were not able to match cases on their characteristics. A dataset of 10,007 individuals with the same date of birth distribution as the CSA live streaming sample, and who had transferred money from Australia to the Philippines between January 2006 and February 2019, were identified by AUSTRAC. The same procedure was then undertaken as with the CSA live streaming group, with the data linked to criminal histories by ACIC, before being provided in a de-identified format to the Australian Institute of Criminology. This process resulted in complete data for a total sample of 209 individuals who had transacted with facilitators of CSA live streaming, and a comparison group of 10,007 individuals, who had sent 288,587 transactions, none of which had been identified as going to facilitators of CSA.

**Figure 1: Distribution of year of birth for those who transacted with facilitators of live streaming of child sexual abuse, and the matched comparison group (% of sample)**




Note: The CSA live streaming group and the comparison group had the same age distribution due to the matching criteria

Source: Philippines CSA live stream financial transaction and comparison dataset

## Measures

This research aimed to develop financial risk indicators identifying those who transact with facilitators of CSA live streaming. Importantly, there were several comparable variables across both groups, including the largest value sent in a single transaction, the smallest value sent in a single transaction, the total number of transactions processed, the average value of all transactions, the total value of all transactions, and the age of the individual at the time of their first transfer. These characteristics were included as independent variables for analysis, with the outcome variable whether an individual had transacted with a known facilitator of CSA live streaming. Each of these characteristics related only to transactions made from Australia to the Philippines, regardless of the payment service used to process the transaction. Every individual included for analysis had made at least one transaction sending money from Australia to the Philippines between January 2006 and February 2019.

## Analytical approach

This research sought to develop a model capable of identifying individuals who were procuring CSA live streaming and individuals who were not, in a sample of more than 10,000 individuals, only using the details of their financial transactions. To do this, we computed a prediction model using machine learning methods, an analytical approach which is particularly good at making accurate predictions using very large and complex datasets. We first produced the machine learning model, known as the random forest, and we then measured how accurately that model performed at predicting which individuals did and did not transact with CSA live streaming facilitators. Next, given the performance of the model, we used post-hoc tests to identify the financial characteristics most useful in making good predictions, and present them as financial risk indicators. Full detail of the modelling procedure follows.

This was a very difficult classification task; the data were significantly imbalanced, with more than 10,000 cases who were not identified as transacting with CSA live streaming facilitators and 209 cases who were. It could have been possible to synthetically inflate the number of cases in the data who had transacted with a CSA live streaming facilitator, but for this research we preferred for the task to be analytically difficult as it better reflected the detection environment—there are many more people in banking data who do not pay to view CSA live streaming than do. To demonstrate the capacity of this modelling procedure to find a 'needle in a haystack', the data were retained as they were available, rather than being synthetically manipulated. This meant that randomisation needed to be effective. As a first step, data were randomised and partitioned into a 70 percent training set and a 30 percent test set. These splits were validated to ensure the distribution of CSA live streaming cases between the groups was correct. The random forest algorithm was trained on the larger set and the model was tested using the partitioned test set (Hyndman & Anthanasopoulos 2014).

A receiver operating characteristic (ROC) curve was used to determine the robustness of the random forest model. A ROC curve plots the true positive rate of classification, referred to as sensitivity (y-axis) compared with the false positive rate, equal to 100 minus the specificity (x-axis) at any threshold value. The area under the receiver operating characteristic (AUROC) curve was then calculated to provide an overall measure of performance, or model accuracy. As a validation exercise, we compared the AUROC for the random forest to the AUROC for a logistic regression model to measure whether the random forest outperformed the logistic regression. We then implemented a bootstrap test for statistical significance between ROC curves to determine whether the differences between the predictive accuracy of the models were statistically significant.

To find the most robust modelling approach, hyperparameters of the random forest model were tuned to optimise iterations and variables randomly considered at each split. When optimising the random forest, model performance will typically plateau as the ntree parameter (number of trees) reaches several hundred iterations (Couronné, Probst & Boulesteix 2018). In the present research, model performance was optimal at ntree=5,000, with five features randomly selected at each split. We then reported the out-of-bag error estimate, which describes the aggregate error of the random forest on the training set (Schonlau & Zou 2020). The random forest is interpreted through the mean decrease Gini coefficient (Hong, Xiaoling & Hua 2016), which provides the proportion of the model accounted for by each variable. The higher the mean decrease Gini, the more important the variable in predicting CSA live streaming. A confusion matrix was then computed to determine the accuracy of the random forest predictions in the test set. The confusion matrix compares the predictions to observed outcomes to measure the performance of the trained model on the test set, providing a metric of how often the model successfully or unsuccessfully made predictions (Barnes & Hyatt 2012).

Finally, post-hoc partial dependence plots (PDPs) were computed for independent variables (Zhao & Hastie 2021) to provide the effect within variables. PDPs show the association between particular characteristics within variables, and CSA live streaming. The y-axis of PDPs represents the logit value associated with different points within the range of a variable, identifying where the relationship with CSA live streaming transactions is strongest (at the highest point in the plots) and weakest (at the lowest point in the plots).

Analysis was performed using the 'randomForest', 'dplyr', 'pRoc', 'pdp' and 'ggplot2' packages of the statistical analysis software R Studio. The model was trained on individuals who procured live streaming of CSA.

### Limitations

While we can be certain that these transactions involved money being sent to known facilitators of CSA live streaming, we cannot be sure that each transaction was intended for that purpose. It is possible that they may have been for other sexual purposes, such as live streaming of adult sexual content. It is, however, unlikely that these transactions were for contact offending or non-sexual purposes, given the receivers of money were arrested for facilitating the sexual exploitation of children (Brown, Napier & Smith 2020). Additionally, the data considered here relate to a single law enforcement operation in the Philippines, making this a type of within-groups analysis. It is unclear whether individuals in this dataset, and therefore the resulting findings, are representative of live streaming offenders more broadly. While those in the comparison group were not detected transacting with facilitators of CSA live streaming, it is possible that they interacted with facilitators who were not detected. Finally, while we matched the comparison data on available variables, there were limitations to our ability to match beyond the three factors used. Additional demographic characteristics relating to CSA live streaming offenders would likely have improved the accuracy of matching, but given the volume of data it is unlikely that this would have meaningfully influenced the findings.

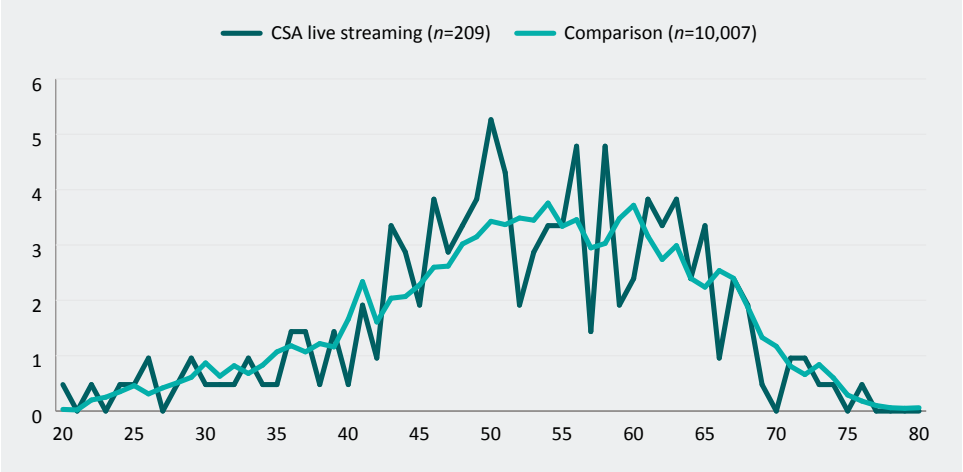
The results presented here are intended as a proof of concept for how financial transactions can be analysed to identify CSA live streaming offenders. However, given the noted limitations, as we progress further from the date of data collection, their usefulness as operationalisable risk indicators may be limited.

## Results

We first compared the available summary statistics for the group who had transacted with a facilitator of CSA live streaming, and the comparison group. Notably, the age distribution of the onset of transactions was similar for both groups (Figure 2). In other words, whether or not these individuals were transmitting money from Australia to the Philippines to procure live streaming of CSA, the proportions of the sample who began transacting at each age were similar.

It is important to note that, due to the matching criteria used to develop these samples, more than half of each group were born between 1937 and 1960, meaning they were between the ages of 46 and 69 years old at the beginning of the observation period. Since the observation period spanned 13 years, these individuals would have been between 59 and 82 years old at the end of it. The matching criteria therefore likely influenced the onset age of transactions among both groups to be older, and the onset opportunity to be within the 13 years from their age in January 2006. However, this does not dictate that the aggregate onset age for the two groups must be similar. It is therefore noteworthy that the average age at which individuals processed the first transaction to a live streaming facilitator was 52.2 years, while among the comparison group the average age of first transaction from Australia to the Philippines was 52.7 years.

**Figure 2: Age of onset of transactions among those who transacted with facilitators of CSA live streaming, and the matched comparison group (%)**




Source: Philippines CSA live stream financial transaction and comparison dataset

There were five transaction characteristics available for analysis in each group (Table 1). The average mean transaction value for procuring CSA live streaming was notably smaller than the average mean transaction value among the comparison group (A$113 vs A$737). On average, the CSA live streaming group made 12 transactions, while the comparison group made 29. The largest (A$236 vs A$2,072) and smallest (A$54 vs A$385) amounts spent in a single transaction tended to be smaller among the CSA live streaming group. Finally, the aggregate value of all transactions tended to be smaller among the CSA live streaming group than the comparison group (A$2,870 vs A$16,143).

**Table 1: Average financial transaction characteristics**

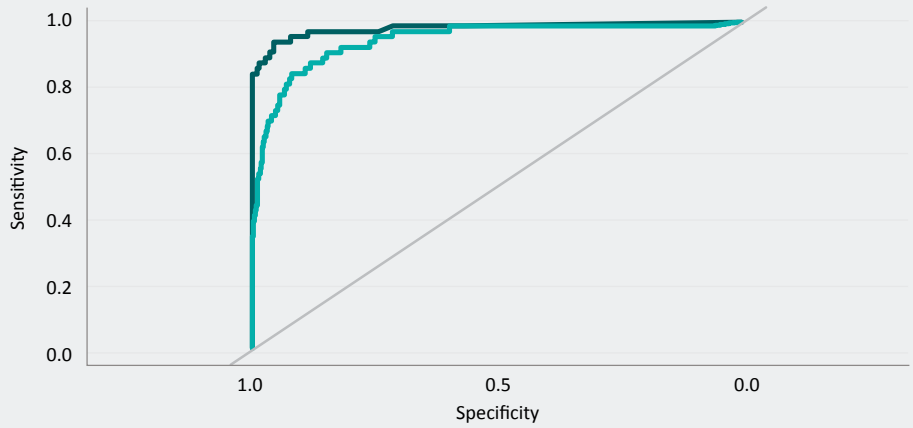
| | CSA live streaming (*n*=209) | Comparison (*n*=10,007) |
|---|---|---|
| Transaction value (A$) (range) | 113 (12–6,745) | 737 (4–211,016) |
| Number of transactions (range) | 12 (1–479) | 29 (1–1,382) |
| Most spent in a single transaction (A$) (range) | 236 (12–9,975) | 2,072 (4–349,991) |
| Least spent in a single transaction (A$) (range) | 54 (1–1,306) | 385 (1–211,016) |
| Collective total of all transactions (A$) (range) | 2,870 (12–119,151) | 16,143 (4–5,330,469) |

Source: Philippines CSA live stream financial transaction and comparison dataset

## Random forest

The AUROC of 0.976 identified that this was a robust model (see Figure 3). The same data were then used to compute a logistic regression, resulting in an AUROC of 0.937. A bootstrap test for statistical significance between ROC curves found that the random forest returned a significantly greater rate of prediction than the logistic regression ($p<0.001$). Of note, we also undertook the same modelling procedure with linked criminal histories for each group included, but the additional data produced only marginally improved accuracy.

**Figure 3: ROC curve for random forest (dark green) and logistic regression (bright green)**




Note: Random forest AUROC=0.976; logistic regression AUROC=0.937

Source: Philippines CSA live stream financial transaction and comparison dataset

The random forest returned an out-of-bag error estimate of 1.3 percent; a confusion matrix was then produced to consider the distribution of prediction accuracy (Table 2). The aggregate classification error closely mirrored the out-of-bag estimate, at 1.3 percent. The rate of false positives was 0.3 percent, while the model had a false negative rate of 46.0 percent. Around one in five individuals (22.7%) identified by the modelling procedure as transacting with facilitators of CSA live streaming were not. However, overall, the model correctly predicted over half (54.0%) of individuals transacting with facilitators.

It is important to note that the sample of CSA live streaming offenders accounted for 2.1 percent of the total sample considered. As a result, if a naïve estimate were to be made that every individual in the total sample did *not* transact with a facilitator of CSA live streaming, it would result in a classification error of 2.06 percent. In other words, such a guess would accurately predict everyone who did not transact with a facilitator of CSA live streaming, and fail to predict everyone who did. It is therefore pivotal that our model significantly outperformed this naïve estimate, producing a lower aggregate classification error of 1.3 percent. While our model accurately predicted which individuals transacted with facilitators of CSA live streaming more than half the time, it almost always correctly predicted which individuals did not transact with facilitators.

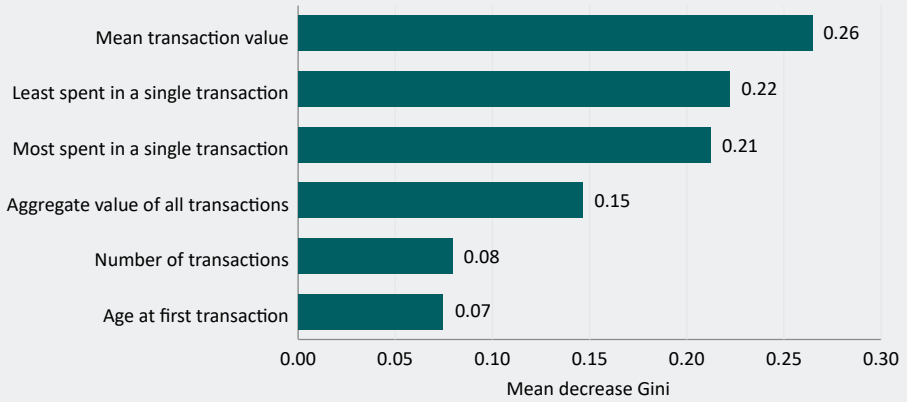
**Table 2: Confusion matrix for random forest model trained on individuals who transacted with facilitators of live streaming of child sexual abuse**

| | | Actual CSA live streaming offender? | | Classification errors |
|---|---|---|---|---|
| | | No | Yes | |
| Predicted CSA live streaming offender (*n*) | No | 2,992 | 29 | 1.0% |
| | Yes | 10 | 34 | 22.7% |
| Classification errors | | 0.3% | 46.0% | 1.3% |

Source: Philippines CSA live stream financial transaction and comparison dataset

Of the six variables used in the modelling process, the five relating to transaction values contributed the greatest power to the predictive model (see Figure 4). The mean transaction value of each individual within these data, the least spent in a single transaction, and the most spent in a single transaction, were the three most important pieces of information in making these predictions. Next most important was the aggregate value of all transactions, while the raw number of transactions made from Australia to the Philippines, and the age of an individual at the time of their first transaction were least important to the predictive power of the model.

**Figure 4: Variable importance associated with live streaming of child sexual abuse**




Source: Philippines CSA live stream financial transaction and comparison dataset

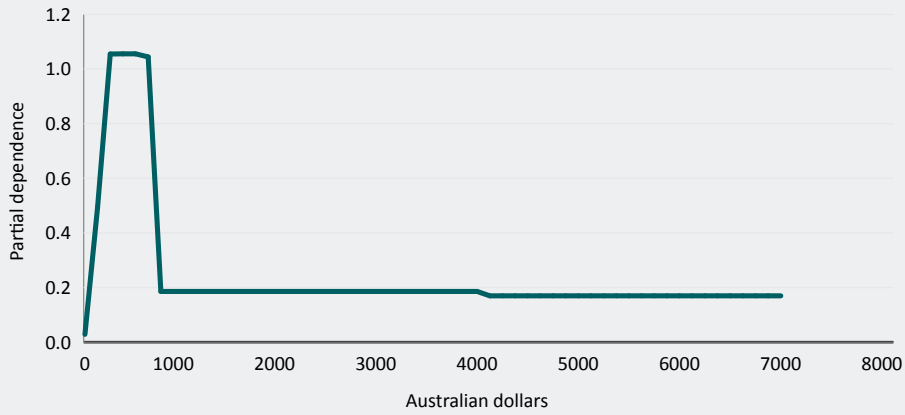
## Partial dependence

We then computed PDPs to find the association with CSA live streaming transactions within each individual variable. In Figures 5 to 10, the point at which the chart peaks demonstrates the strongest relationship with CSA live streaming transactions, or the point of greatest risk, and the lowest point of each chart represents the point at which the interaction is weakest.

### *Average value of transactions*

The average transaction value was the most important variable in the classification rate produced by the random forest. That is, it contributed the largest amount of predictive power in identifying who did and who did not transact with CSA live streaming facilitators. A PDP was computed to identify the point within this variable most associated with individuals who transacted with CSA live streaming facilitators. Figure 5 suggested that individuals with an average transaction value of between A$300 and A$600 across all of their transactions were most associated with CSA live streaming.

**Figure 5: Average transaction value associated with live streaming of child sexual abuse (A$)**

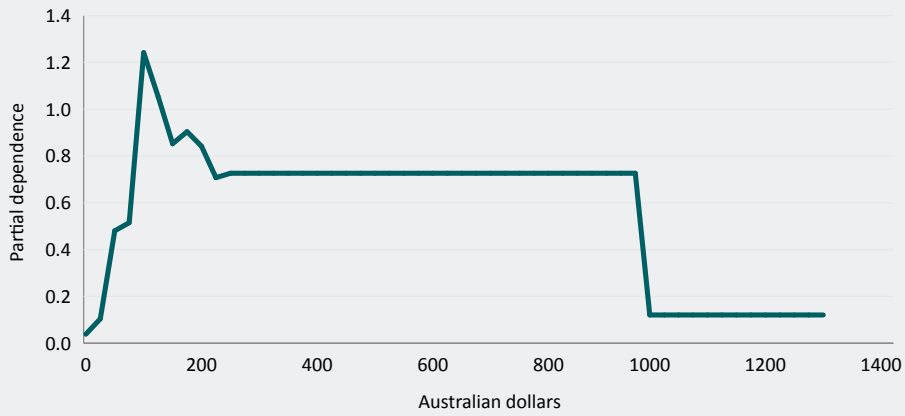



Source: Philippines CSA live stream financial transaction and comparison dataset

### *The smallest and largest transactions*

For each individual in these data, we analysed the smallest and largest transactions. Figure 6 reflects the relationship between smallest amount sent and whether an individual transacted with a facilitator of CSA live streaming. The smallest transaction having a value between A$50 and A$200 held the strongest association with an individual transacting with a facilitator, but it is important to note that this variable featured a long tail, with some degree of relationship up to A$950.

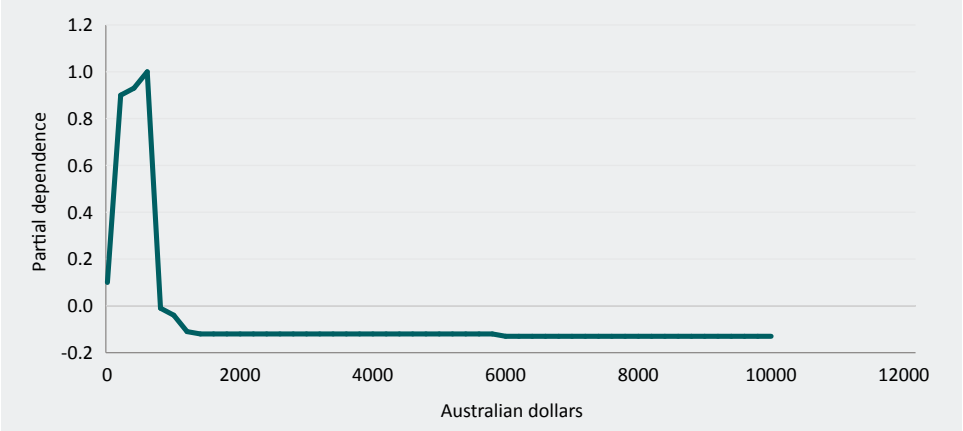
**Figure 6: Least spent in a single transaction associated with live streaming of child sexual abuse (A$)**




Source: Philippines CSA live stream financial transaction and comparison dataset

Taking the value of the largest amount sent to the Philippines by each individual in the data, we then computed a PDP investigating its association with transacting with facilitators of CSA live streaming (Figure 7). Compared with the findings for the minimum spent, this variable indicated a smaller range with an identifiable association. The relationship was strongest when the largest transaction value was between A$200 and A$600. Importantly, after this point the relationship was inverse, meaning that those who made at least one transaction worth at least A$1,000 were less likely to be procuring CSA live streaming.

**Figure 7: Most spent in a single transaction associated with live streaming of child sexual abuse (A$)**

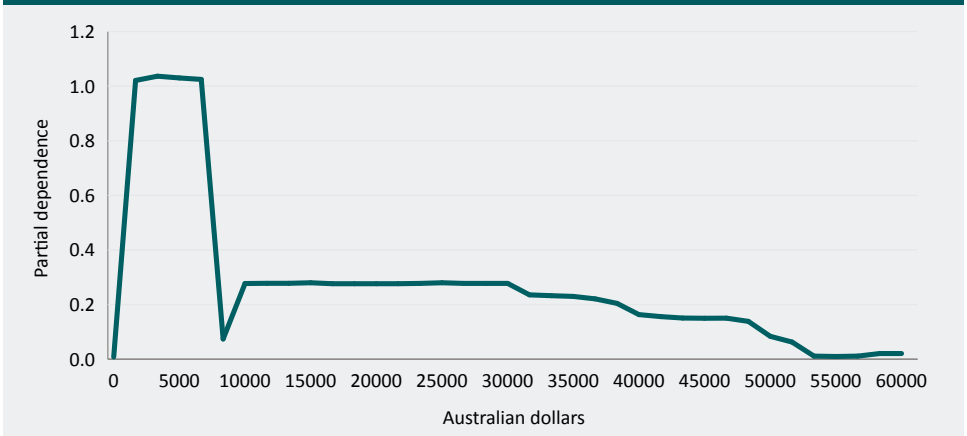



Source: Philippines CSA live stream financial transaction and comparison dataset

### *Total value of all transactions*

A PDP was then produced reflecting the relationship between the total value of all transactions made by each individual in these data, and transacting with a facilitator of CSA live streaming (Figure 8). This relationship was strongest among individuals who had, in total, sent between A$2,000 and A$6,000 from Australia to the Philippines. Importantly, among those who sent more than A$10,000, the relationship was weak; these individuals were unlikely to be transacting with facilitators of CSA live streaming.

**Figure 8: Total amount of money sent associated with live streaming of child sexual abuse (A$)**

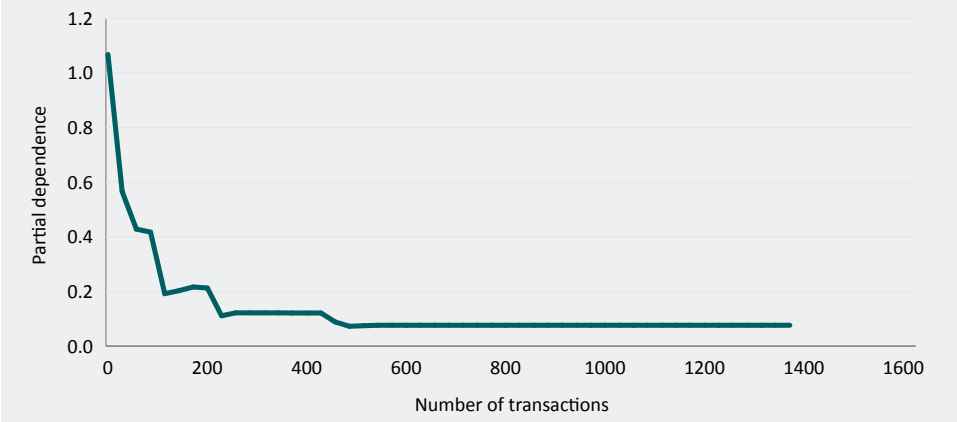



Source: Philippines CSA live stream financial transaction and comparison dataset

### *Number of transactions*

The relationship between the raw number of transactions made and transacting with a facilitator of CSA live streaming was then considered. Figure 9 suggests that individuals who made fewer transactions were more likely to be transacting with a facilitator of CSA live streaming. In particular, those who processed fewer than 30 transactions were most likely to do so, and this interaction diminished as the number of transactions increased. Ultimately, those who had sent money from Australia to the Philippines more than 90 times were notably less likely to be transacting with a facilitator of CSA live streaming.

**Figure 9: Number of transactions associated with live streaming of child sexual abuse**

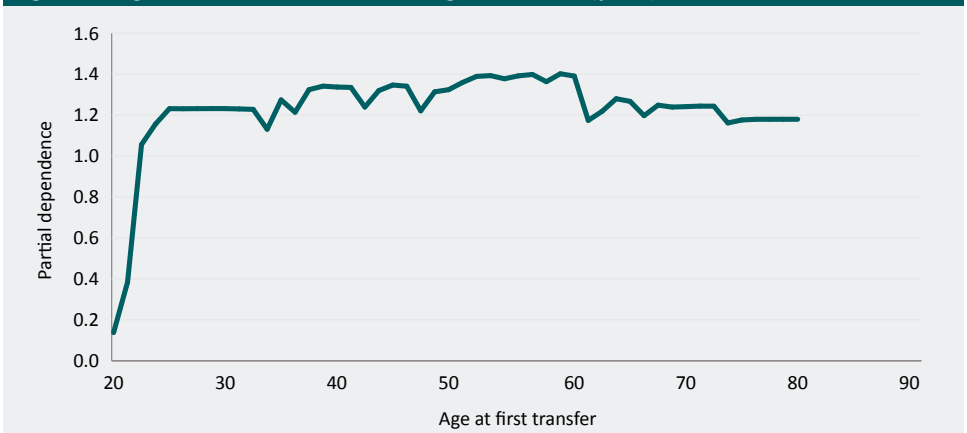



Source: Philippines CSA live stream financial transaction and comparison dataset

### *Age at the time of first transaction*

The final variable included in the model was the age at which an individual made their first transaction. This variable contributed the least power to the classification model and, on producing the PDP in Figure 10, it was clear this was because a range of ages held a relationship with transacting with a facilitator of CSA live streaming. However, it was clear that people under the age of 25 were unlikely to be transacting with facilitators, and those most likely were between the ages of 50 and 60.

**Figure 10: Age at onset of CSA live streaming transactions (years)**




Source: Philippines CSA live stream financial transaction and comparison dataset

## Discussion

This study undertook a 'needle in a haystack' task, using a machine learning approach to develop a proof of concept model for identifying CSA live streaming offenders, with some success. Overall, the model correctly classified 54.0 percent of Australian individuals transacting with CSA live streaming facilitators in the Philippines. Importantly, this was achieved with a relatively low rate of false positives. Indeed, it was estimated that about one in 300 individuals (0.3%) identified by the model as an offender would turn out to be a non-offender.

This research isolated a series of characteristics indicative of an individual transacting with a facilitator of CSA live streaming. Importantly, this type of modelling does not intend to reflect the patterns within the financial transactions used to procure CSA live streaming as explored by Brown, Napier and Smith (2020). Rather, it finds the factors that collectively predict the largest proportion of CSA live streaming offenders possible. Table 3 consolidates these findings into financial risk indicators. None of these indicators should function independently of the others, but collectively they indicate a high probability of CSA live streaming. Our analysis suggests that individuals who send money from accounts in Australia to the Philippines and whose transactions meet the specifications in Table 3 represent a risk of procuring CSA live streaming.

| Table 3: Financial risk indicators for individuals transacting with facilitators of child sexual abuse live streaming |
|---|
| Average transaction value per individual of less than A$600 |
| Smallest transaction value per individual between A$50 and A$200 |
| Largest transaction value per individual between A$200 and A$600 |
| Very unlikely to make individual transactions worth A$1,000 or more |
| Total value of transactions per individual between A$2,000 and A$6,000 |
| Very likely to process fewer than 30 transactions |
| Very unlikely to be 25 or under at the time of first transaction, most likely to be between the ages of 50 and 65 |

Note: 'Transaction' refers to money sent from Australia to the Philippines

It is important to reiterate that, when the factors listed in Table 3 co-occur, the likelihood of CSA live streaming is considerable. For example, the average transaction value for live streaming offenders is known to be relatively low (Brown, Napier & Smith 2020), a finding also reflected in Table 1 and Figure 4. As a standalone measure, average transaction values could only assist to predict a small number of CSA live streaming cases (see Figure 3). However, when cases were assessed as meeting the sequence of indicators in Table 3 in totality, the likelihood of successfully predicting CSA live streaming was high. Furthermore, intelligence-based law enforcement reports may yield individual indicators for CSA live streaming transactions that differ to those presented here, given the present type of modelling is rarely used in this context. However, AUSTRAC (2022) reported that CSA live streaming transactions are usually less than A$500, which is similar to the average transaction value presented in Table 3.

### Implications and limitations

As a proof of concept, this model performed exceptionally well, providing evidence that it may be possible to predict CSA live streaming offenders using financial transactions data. It is possible that, in collaboration with law enforcement and regulatory agencies, a model such as this could be developed to detect suspicious transactions for further investigation. When cases met the risk criteria, investigations could be supported by existing financial intelligence holdings and, where appropriate, could lead to downstream intervention with offenders. This type of approach could also be used for real-time monitoring of suspicious transactions, assisting in identification, referral and investigation procedures. The performance of this proof of concept model supports consideration as an approach to assist the detection and triage of cases.

However, there are also several limitations to this research. While the model demonstrated success in making predictions, the data were current up to 2019. It is possible that the CSA live streaming environment has changed in the intervening time, meaning our model is less useful. For this reason, we suggest that it is a proof of concept rather than an implementable procedure. If this model were to be implemented, it must first be either validated or reproduced with contemporary data.

Finally, while our research focuses on the Philippines, CSA live streaming is not restricted solely to this region. Rather, any country which experiences elevated rates of poverty coinciding with high-speed internet appears to be vulnerable to facilitation of CSA live streaming. It is possible that the characteristics of these offences occurring in other parts of the world could be different to those we have observed. Separately, the increasing potential use of cryptocurrency to procure CSA live streaming offences is noteworthy. We do not yet know if the financial values observed in this study translate to instances where cryptocurrency is used to procure offences. While we are unable to make this assessment using the present data, future research should seek to understand the role of cryptocurrency in supporting these offences, and whether the type of currency influences the trends in procuring offences.

## Conclusion

Risk indicators may hold important potential as a tool for identifying high-risk individuals and cases which need further investigation to determine if CSA live streaming has occurred. Our findings suggest this may be possible, with very few false positives—avoiding misidentification of those who are not involved in this type of offending. This proof of concept model offers support for an approach to targeting individuals who have previously operated with a significant degree of impunity, ultimately reducing some of the harm caused to children through live streamed CSA.

## Acknowledgements

The Australian Institute of Criminology would like to acknowledge AUSTRAC for partnering with us on the project and providing transaction data.

## References

*URLs correct as at March 2025*

Açar KV 2017. Webcam child prostitution: An exploration of current and futuristic methods of detection. *International Journal of Cyber Criminology* 11(1): 98–109. https://doi.org/10.5281/zenodo.495775

AUSTRAC—see Australian Transaction Reports and Analysis Centre

Australian Transaction Reports and Analysis Centre 2022. *Combating the sexual exploitation of children for financial gain: Financial crime guide.* AUSTRAC. https://www.austrac.gov.au/business/how-comply-guidance-and-resources/guidance-resources/combating-sexual-exploitation-children-financial-gain

Australian Transaction Reports and Analysis Centre 2019. *Combatting the sexual exploitation of children for financial gain: Activity indicators*. AUSTRAC. https://www.austrac.gov.au/business/how-comply-guidance-and-resources/guidance-resources/combating-sexual-exploitation-children-financial-gain-activity-indicators-report

Barnes GC & Hyatt JM 2012. *Classifying adult probationers by forecasting future offending.* https://nij.ojp.gov/library/publications/classifying-adult-probationers-forecasting-future-offending

Batha E 2016. Internet providers urged to tackle live-streaming of child sex in the Philippines. Thomas Reuters Foundation, 7 June. https://www.reuters.com/article/us-philippines-internet-child-sexcrimes/internetproviders-urged-totackle-live-streaming-of-child-sex-in-the-philippines-idUSKCN0YT04W

Brown R, Napier S & Smith RG 2020. Australians who view live streaming of child sexual abuse: An analysis of financial transactions. *Trends & issues in crime and criminal justice* no. 589. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti04336

Couronné R, Probst P & Boulesteix AL 2018. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics* 19(270): 1–14. https://doi.org/10.1186/s12859-018-2264-5

Cubitt T, Napier S & Brown R 2023. Understanding the offline criminal behaviour of individuals who live stream child sexual abuse. *Journal of Interpersonal Violence* 38 (9–10): 6624–6649. https://doi.org/10.1177/08862605221137712

Cubitt T, Napier S & Brown R 2021. Predicting prolific live streaming of child sexual abuse. *Trends & issues in crime and criminal justice* no. 634. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78320

DeHart D, Dwyer G, Seto MC, Moran R, Letourneau E & Schwarz-Watts D 2017. Internet sexual solicitation of children: A proposed typology of offenders based on their chats, e-mails, and social network posts. *Journal of Sexual Aggression* 23(1): 77–89. https://doi.org/10.1080/13552600.2016.1241309

ECPAT International 2017. Online child sexual exploitation: An analysis of emerging and selected issues. *ECPAT International Journal* 12: 1–63

EFC—see European Financial Coalition against Commercial Sexual Exploitation of Children Online

European Financial Coalition against Commercial Sexual Exploitation of Children Online (EFC) 2015. *Strategic assessment 2014*. https://www.europol.europa.eu/publications-documents/commercial-sexual-exploitation-of-children-online

Europol 2020a. *Exploiting isolation: Offenders and victims of online child sexual abuse during the COVID-19 pandemic*. https://www.europol.europa.eu/publications-documents/exploiting-isolation-offenders-and-victims-of-online-child-sexual-abuse-during-covid-19-pandemic

Europol 2020b. *IOCTA 2020: Internet organised crime threat assessment*. Luxembourg: Publications Office of the European Union. https://www.europol.europa.eu/publications-events/main-reports/internet-organised-crime-threat-assessment-iocta-2020

Europol 2019. *IOCTA 2019: Internet organised crime threat assessment.* Luxembourg: Publications Office of the European Union. https://www.europol.europa.eu/publications-events/main-reports/internet-organised-crime-threat-assessment-iocta-2019

Financial Action Task Force 2025. *Detecting, disrupting and investigating online child sexual exploitation: Using financial intelligence to protect children from harm.* Paris: Financial Action Task Force. https://www.fatf-gafi.org/en/publications/Fatfgeneral/Online-child-sexual-exploitation.html

Gibraltar Financial Intelligence Unit nd. *Combating online child sexual abuse & exploitation: Guidance notes*. https://www.gfiu.gov.gi/what-is-online-child-sexual-abuse-and-exploitation

Global Alliance against Child Sexual Abuse Online 2016. *2015 Threat assessment report*. https://home-affairs.ec.europa.eu/system/files/2020-09/global_alliance_threat_assessment_en.pdf

Hong H, Xiaoling G & Hua Y 2016. *Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on random forest*. Paper to 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 26–28 August. https://ieeexplore.ieee.org/document/7883053

Hyndman RJ & Anthanasopoulos G 2014. *Forecasting: Principles and practice*. Melbourne: Monash University

International Justice Mission 2023. *Scale of harm: Research method, findings and recommendations*. https://www.ijmuk.org/stories/1-in-100-children-sexually-exploited-in-livestreams-new-abuse-images-and-videos-in-the-philippines-last-year-driven-by-foreign-demand

Masri L 2015. *Webcam child sex abuse*. CUNY Academic Works. https://academicworks.cuny.edu/gj_etds/64

Napier S, Teunissen C & Boxall H 2021. How do child sexual abuse live streaming offenders access victims? *Trends & issues in crime and criminal justice* no. 642. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti78474

Napier S, Seto MC, Cashmore J & Shackel R 2024. Characteristics that predict exposure to and subsequent intentional viewing of child sexual abuse material among a community sample of internet users. *Child Abuse & Neglect* 156: 106977. https://doi.org/10.1016/j.chiabu.2024.106977

Netclean 2019. *Netclean report 2019: A report about child sexual abuse crime*. Gothenberg: Netclean. https://respect.international/netclean-report-2019/

Puffer E, McDonald K, Pross M & Hudson D 2014. *Webcam child sex tourism: An emerging global issue.* https://digitalcommons.cedarville.edu/research_scholarship_symposium/2014/podium_presentations/15/

Ramiro LS, Martinez AB, Tan JRD, Mariano K, Miranda GMJ & Bautista G 2019. Online child sexual exploitation and abuse: A community diagnosis using the social norms theory. *Child Abuse and Neglect* 96: 104080. https://doi.org/10.1016/j.chiabu.2019.104080

Schonlau M & Zou RY 2020. The random forest algorithm for statistical learning. *Stata Journal: Promoting communications on statistics and Stata* 20(1): 3–29. https://doi.org/10.1177/1536867X20909688

Terre des Hommes 2013. *Webcam child sex tourism: Becoming Sweetie: A novel approach to stopping the global rise of webcam child sex tourism*. The Hague: Terre des Hommes. https://www.tdh.org/en/digital-library/documents/webcam-child-sex-tourism

Zhao Q & Hastie T 2021. Causal interpretations of black box models. *Journal of Business & Economic Statistics* 37(1): 272–281. https://doi.org/10.1080/07350015.2019.1624293

This work was originally published as *Trends & issues in crime and criminal justice* no. 718. It is available online at **https://doi.org/10.52922/ti77956**.

# 10. Sexual extortion of Australian adolescents: Results from a national survey

Heather Wolbers, Timothy Cubitt, Sarah Napier, Michael John Cahill, Mariesa Nicholas, Melanie Burton and Katherine Giunta

## Introduction

Sexual extortion is a form of blackmail in which a perpetrator threatens to release intimate material of a victim unless they comply with certain demands (Açar 2016; Alsoubai et al. 2022; Liggett 2019; Patchin & Hinduja 2020; Wolak et al. 2018; Wolbers et al. 2025). Sexual extortion perpetrators typically demand money, additional intimate materials or sexual contact with the victim (Patchin & Hinduja 2020; Wolak et al. 2018). The eSafety Commissioner (2023a) received a significant increase in reports of sexual extortion in the first quarter of 2023, while the US National Center for Missing and Exploited Children (NCMEC) reported a 150 percent increase in such reports from children globally from 2022 to 2023 (NCMEC 2024).

Research suggests that individuals going through adolescence often want to achieve a clearer sense of identity (Branje et al. 2021), which can co-occur with more risk-taking and sensation-seeking (Duell et al. 2018). The integration of the internet into young people's lives means that, increasingly, these developmental processes are occurring alongside the introduction of the digital landscape and online domains (Crone & Konijn 2018). The combination of these factors can result in increases in risky sexual behaviour online among adolescents, such as sharing intimate images with someone they have met in person or only know online (Madigan et al. 2018), which increases vulnerability to perpetrators of sexual extortion (Patchin & Hinduja 2024).

The impacts of sexual extortion can be diverse and severe. A systematic review of 12 studies found that experiences of sexual extortion can have adverse mental health and wellbeing outcomes for children, ranging from anxiety and shame to self-harm and suicide (Wolbers et al. 2025). The study also identified negative social impacts such as losing friends or feeling isolated, and negative impacts on victims' work, school and home life. A survey of 18- to 25-year-olds in the United States revealed that sexual extortion can sometimes be a prolonged experience, particularly if the victim was a minor at the time of the incident (Wolak et al. 2018), which may exacerbate the impacts on victims' wellbeing. Further, as cases of sexual extortion involve the production and distribution of intimate or sexual material, when the victim is a minor, incidents can contribute to the proliferation of child sexual abuse material online (eSafety Commissioner 2023b), and have associated harms for victims (Gewirtz-Meydan et al. 2018).

## Experiences of sexual extortion

There can be considerable variation in how a young sexual extortion victim meets and interacts with a perpetrator and how they experience sexual extortion (Patchin & Hinduja 2020; Wolak et al. 2018). Importantly, research into the nature and experiences of sexual extortion among young people tends to report mixed findings. Some survey-based studies indicate that sexual extortion experienced by minors is most commonly perpetrated by a current or former intimate partner (Finkelhor, Turner & Colburn 2023, 2022; Wolak et al. 2018), while other non-survey studies (ie using data from tiplines, court documents or online posts) suggest perpetrators are more likely to be someone met online (Alsoubai et al. 2022; Australian Centre to Counter Child Exploitation (ACCCE) 2022; Canadian Centre for Child Protection (C3P) 2024; O'Malley & Holt 2022).

Differences across the literature can be attributed to diverse data sources or definitions and measures of sexual extortion. For example, a community survey of 5,568 young people aged 12 to 17 years in the United States found boys were more likely to experience sexual extortion than girls (5.8% vs 4.1%; Patchin & Hinduja 2020). Similarly, other research examining reports to tiplines found most reports were from boys (90–100%; ACCCE 2022; C3P 2024; Cross et al. 2022). Surveys of sexual extortion victims, however, tend to find victims are primarily girls (77–91%; Finkelhor, Turner & Colburn 2023, 2022; Wolak et al. 2018).

In addition to the mixed findings on the sociodemographic characteristics of victims and common circumstances of sexual extortion, the ways in which sexual extortion is perpetrated may also be changing. For example, emerging evidence suggests the financial gains from sexual extortion have attracted the attention of organised crime groups (Raffile et al. 2024; Thorn & NCMEC 2024). These groups may be particularly targeting young boys in large numbers to blackmail them for money (ACCCE 2022; C3P 2024; Thorn & NCMEC 2024).

Due to the constant evolution of technologies including artificial intelligence (Internet Watch Foundation 2023), and the ability for offenders to communicate with and learn from each other online (Raffile et al. 2024), changes in the nature of sexual extortion could occur rapidly. It is therefore important to better understand the scope and nature of sexual extortion of young people, to improve capacity for design and implementation of prevention and disruption approaches. The current study examines the prevalence and nature of sexual extortion among adolescents in Australia.

# Method

## Data collection and sample

This study uses data from a survey of adolescents between the ages of 16 and 18 in Australia, examining experiences of sharing sexual material, catfishing and sexual extortion. Conducted by Octopus Group Surveys in mid-2024, the survey was sent to panels of individuals who had voluntarily joined to receive incentives in exchange for completing surveys. Participants were recruited from the Octopus, Clint, PureSpectrum and Student Edge panels. Proportional quota sampling was initially used to ensure the sample was reflective of the spread of adolescents in Australia. Quotas were stratified by age, sex and usual place of residence, based on Australian Bureau of Statistics (2023) population data.

The raw completion rate for invitations sent to this panel was 4.8 percent, many of whom were ineligible because they or their children were not within the target age range of the survey. However, 23.9 percent of those eligible completed the survey. Given the sensitive subject matter of the survey, in combination with the target age group, this was a strong rate of completion. Respondents completed the survey, on average, within 11 minutes. Respondents who provided nonsensical responses or had unusually quick completion times were removed from the sample. A subsample of respondents initially contacted via their parents were also removed due to inconsistencies in the sampling procedure. Compared with the distribution created by proportional quota sampling, this resulted in slightly higher proportions of women (55.6% vs 48.7%) and 18-year-olds (44.4% vs 35.4%) and slightly lower proportions of men (41.5% vs 48.7%) and 16- (22.0% vs 28.7%) and 17-year-olds (33.6% vs 36.0%). The geographical distribution of respondents did not change notably. The findings of this research reflect the sample of 1,953 respondents who were recruited directly.

## Measures and analytical approach

The focus of this study was to gain a better understanding of adolescents' experiences of sexual extortion. To measure sexual extortion, survey respondents were asked: 'Has anyone ever demanded something from you (eg money or photos) and threatened to post or share an intimate or nude image/video of you if you didn't do what they say?' Respondents who had experienced sexual extortion were then asked follow-up questions about their experiences, focusing on the recency and frequency of the extortion, age of first experience, threats and demands received, and whether they had met the perpetrator online. We then asked more specific questions about the online platform(s) on which they met the perpetrator, whether the perpetrator asked them to switch to other platforms, and whether they complied with perpetrator demands. First, we examined the characteristics of respondents and experiences of those who were sexually extorted in their lifetime. We then focused on whether experiences differed according to the gender of respondents and whether they had met the perpetrator online.

## Sample characteristics

Of the full sample, 41.5 percent identified as men, 55.6 percent as women and 2.6 percent as trans or gender diverse (Table 1). The mean age of respondents was 17.2 (*SD*=0.8), with 22.0 percent being 16 years of age, 33.6 percent 17 years of age and 44.4 percent 18 years of age. Further, 15.4 percent of respondents identified as lesbian, gay, bisexual, unsure/questioning or another non-heterosexual sexual identity. One in 20 (5.5%) were Aboriginal and/or Torres Strait Islander and 9.7 percent were from a non-English-speaking background. Most respondents were studying in high school or university (84.6%), while 10.0 percent were working and 4.7 percent were neither working nor studying.

| Table 1: Sample characteristics (*n*=1,953) | | |
|---|---|---|
| | *n* | % |
| **Gender** | | |
| Men | 811 | 41.5 |
| Women | 1,085 | 55.6 |
| Trans and gender diverse[a] | 50 | 2.6 |
| **Age** | | |
| 16 | 430 | 22.0 |
| 17 | 656 | 33.6 |
| 18 | 867 | 44.4 |
| **Sexually diverse[b]** | 300 | 15.4 |
| **Aboriginal and/or Torres Strait Islander** | 107 | 5.5 |
| **Non-English-speaking background** | 190 | 9.7 |
| **Long-term health condition[c]** | 422 | 21.6 |
| **Daily activity** | | |
| Studying | 1,652 | 84.6 |
| Working | 196 | 10.0 |
| Not studying or working | 91 | 4.7 |

a: Includes respondents who identified as trans men, trans women, non-binary, unsure/questioning, gender fluid or gender queer

b: Lesbian, gay, bisexual, unsure/questioning or another non-heterosexual sexual identity

c: Includes any health condition that has lasted, or is likely to last, 6 months or longer

Note: Denominators include respondents who were unsure or did not disclose information

Source: AIC survey of adolescents, 2024 [computer file]

## Limitations

This research relied on data collected by surveying adolescent members of online panels; thus, we are cautious to not generalise the findings across the wider community or beyond this age range. Our study also considers sexual extortion on an individual level rather than an incident level, meaning for adolescents who were sexually extorted more than once we are unable to determine specific details of each incident or differences between incidents.

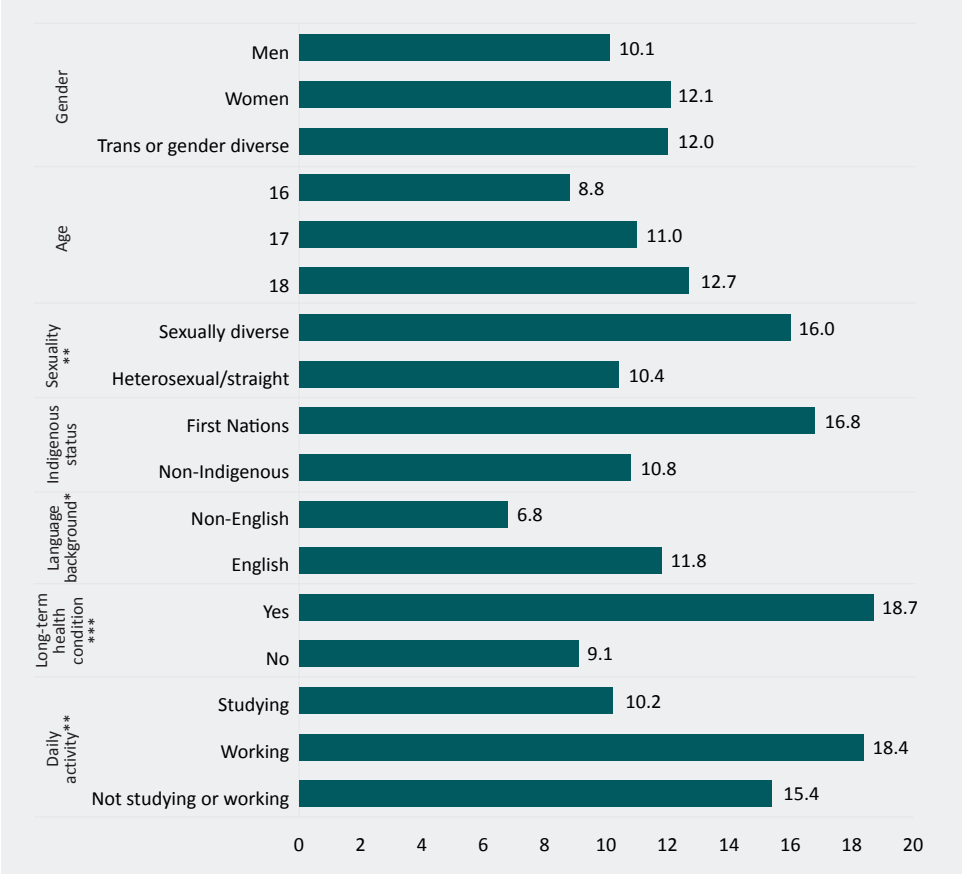
# Results

Of the full sample (*n*=1,953), 220 respondents or 11.3 percent indicated they had experienced sexual extortion in their lifetime (*n*=18 did not disclose whether they had or had not).

## Vulnerability to sexual extortion

We first examined the prevalence of sexual extortion aggregated by respondent sociodemographic characteristics and used chi-square analyses to test for associations (Figure 1). Gender was not significantly associated with sexual extortion experiences, with women (12.1%), men (10.1%) and trans or gender diverse respondents (12.0%) experiencing fairly similar rates. Each consecutive age reported experiencing higher rates of sexual extortion, but this was likely a function of opportunity, with 18-year-olds (12.7%) having had more time to experience sexual extortion than 17-year-olds (11.0%) or 16-year-olds (8.8%). Adolescents who identified as sexually diverse (16.0% vs 10.4%, $\chi^2(1)=8.2$, $p<0.01$), had a long-term health condition (18.7% vs 9.1%, $\chi^2(1)=30.9$, $p<0.001$), and came from an English-speaking background (11.8% vs 6.8%, $\chi^2(1)=4.2$, $p<0.05$) reported significantly higher rates of sexual extortion. Finally, adolescents who were working (18.4%), or neither working nor studying (15.4%), reported significantly higher rates of sexual extortion than those who were studying (10.2%, $\chi^2(2)=13.6$, $p<0.01$).

**Figure 1: Prevalence of sexual extortion, by sociodemographic characteristics (%) (*n*=1,953)**



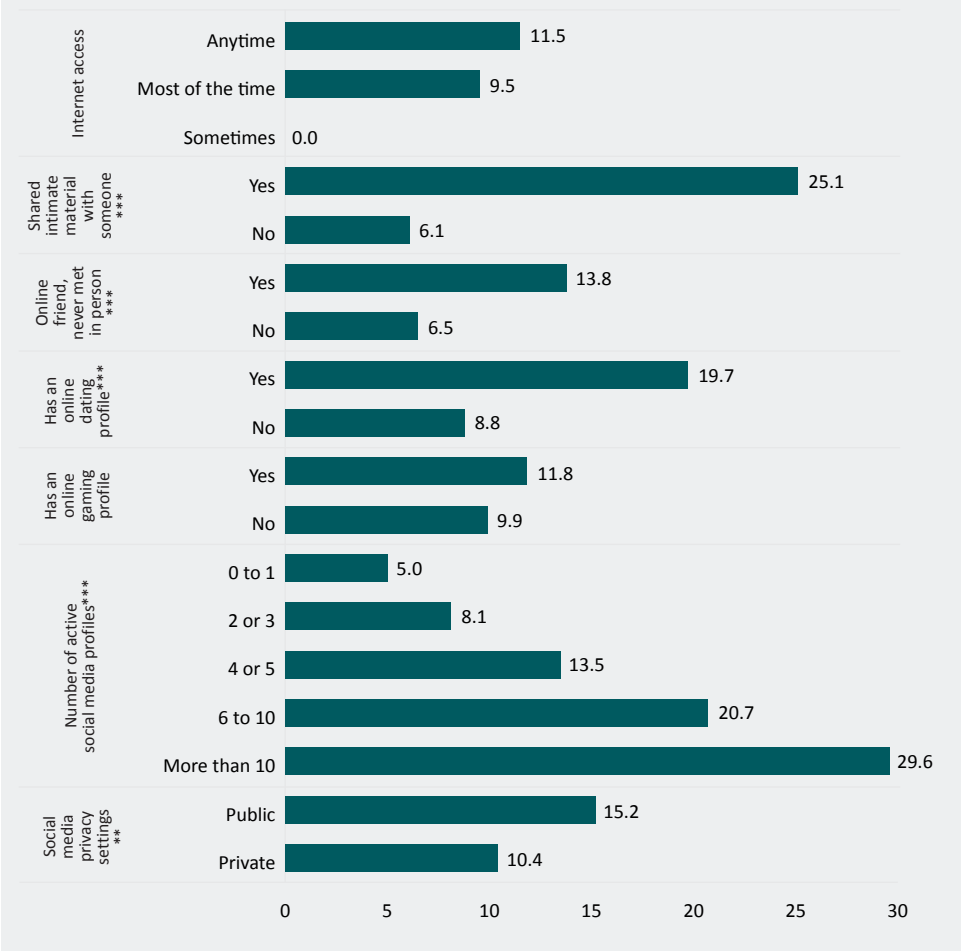

***statistically significant at $p<0.001$, **statistically significant at $p<0.01$, *statistically significant at $p<0.05$

Note: Denominators include respondents who did not disclose whether they had experienced sexual extortion. Missing data were excluded from chi-square tests. We did not test for an association between sexual extortion and respondent age, as any observed difference would primarily be the product of opportunity effects and would not be meaningful

Source: AIC survey of adolescents, 2024 [computer file]

When examining the online activities of adolescents, we observed a significantly higher prevalence of sexual extortion among adolescents who had ever shared intimate material with someone (25.1% vs 6.1%, $\chi^2(1)=135.8$, $p<0.001$), who had an online friend they had never met in person (13.8% vs 6.5%, $\chi^2(1)=22.9$, $p<0.001$), and who had made an online dating profile (19.7% vs 8.8%, $\chi^2(1)=42.1$, $p<0.001$) but not a gaming profile (Figure 2). Sexual extortion was significantly more common among those who used a larger number of social media platforms ($\chi^2(4)=42.8$, $p<0.001$), peaking at 29.6 percent among those on more than 10 social media platforms. Finally, sexual extortion was significantly less prevalent among adolescents who typically set their social media accounts to private rather than public (10.4% vs 15.2%, $\chi^2(1)=7.4$, $p<0.01$). Level of internet access was not significantly associated with sexual extortion.

**Figure 2: Prevalence of sexual extortion, by online activities (%) (*n*=1,953)**




***statistically significant at $p<0.001$, **statistically significant at $p<0.01$

Note: Denominators include respondents who did not disclose information. Missing data were excluded from chi-square tests. Adolescents who accessed the internet 'sometimes' were excluded from chi-square analysis because none had experienced sexual extortion

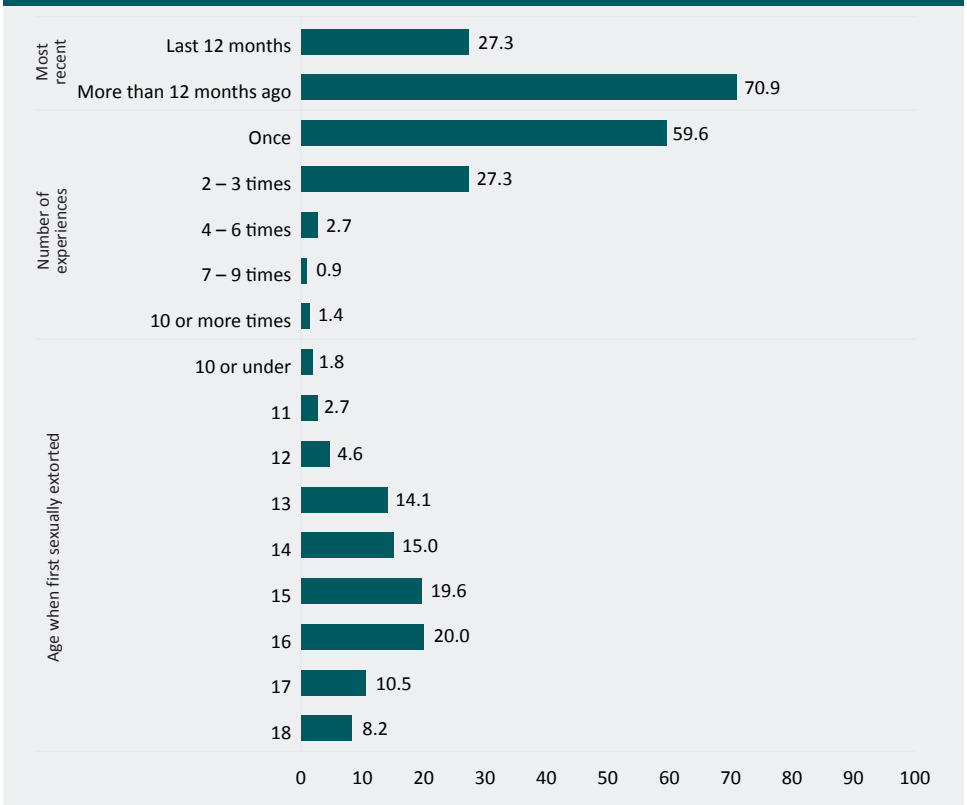
Source: AIC survey of adolescents, 2024 [computer file]

## The nature of sexual extortion

### *Age of victims and frequency of sexual extortion*

Among adolescents who had been sexually extorted (*n*=220), 27.3 percent said it happened in the last 12 months (Figure 3). While most experienced one event (59.6%), one-third experienced two or more incidents (32.3%). Of this group, the majority experienced two or three sexual extortion events (27.3%), while 5.0 percent experienced sexual extortion more than three times. Of the 220 victims, 88.2 percent (*n*=194) were under the age of 18 when they were first sexually extorted, while a minority were 18 years of age (8.2%, *n*=18), or did not report their age at the first incident (3.6%, *n*=8). Over half of victims had been sexually extorted for the first time when under the age of 16 (57.7%), while one in 10 were under the age of 13 (9.1%). The most common age at which adolescents first experienced sexual extortion was 15 or 16 (19.6% and 20.0% respectively).

**Figure 3: Age and frequency of sexual extortion among adolescents (%) (*n*=220)**



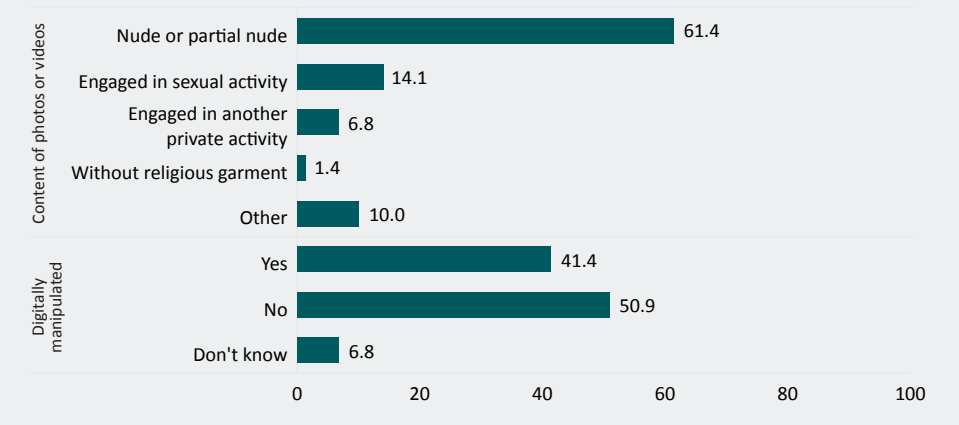

Note: Denominators include respondents who did not disclose information

Source: AIC survey of adolescents, 2024 [computer file]

### *Types of material used to sexually extort victims*

The most common type of content used to perpetrate sexual extortion were photos or videos depicting adolescents as nude or partially nude (61.4%; Figure 4), followed by material depicting them engaged in sexual activity (14.1%) or in another private activity (eg using the toilet; 6.8%), and material depicting the adolescent without their religious garment (1.4%). Two in five adolescents who were sexually extorted said the extortion material had been digitally manipulated (41.4%).

**Figure 4: The nature of the material used to sexually extort adolescents (%) (*n*=220)**



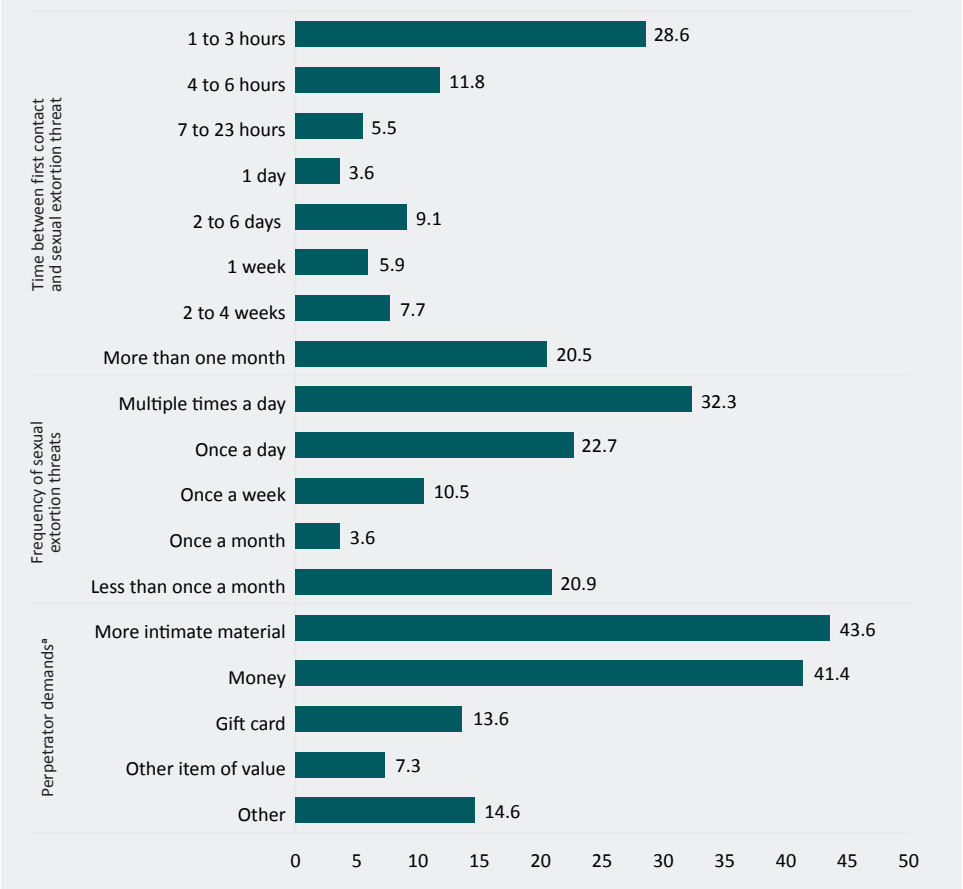

Note: Denominators include respondents who did not disclose information

Source: AIC survey of adolescents, 2024 [computer file]

### *Characteristics of threats and demands*

Most respondents said that the time between first contact with a perpetrator and the first sexual extortion threat was less than a week (58.6%; Figure 5), while just under half (45.9%) said it was less than one day. Just under a third of adolescents were threatened multiple times a day (32.3%); however, notable proportions were threatened either monthly or less (24.5%) or daily (22.7%), suggesting variation in experiences. The most common demand was for more intimate materials (43.6%), followed by money (41.4%), a gift card (13.6%) or something else of financial value (7.3%). Half of victims received a financial demand for money, a gift card or something else of financial value (50.5%). Around one in 10 adolescents received a financial demand and a demand for intimate material (9.6%; Figure 6).

**Figure 5: The nature of sexual extortion threats and demands received by adolescents (%) (*n*=220)**

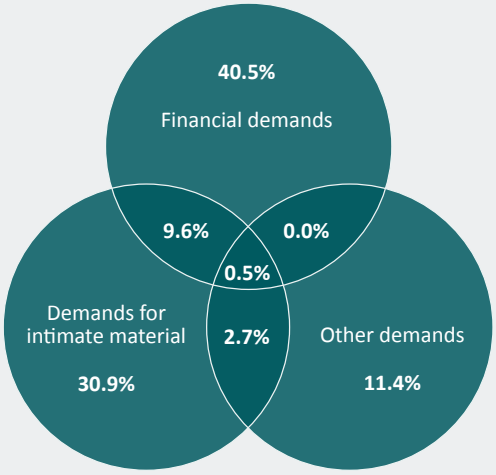



a: Respondents could select multiple options for this survey question

Note: Denominators include respondents who did not disclose information

Source: AIC survey of adolescents, 2024 [computer file]

**Figure 6: The overlap of the types of sexual extortion demands received by adolescents (*n*=220)**



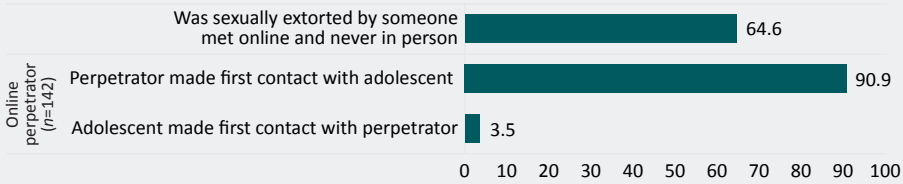

Note: Denominators include 10 respondents who did not disclose demands received

Source: AIC survey of adolescents, 2024 [computer file]

### *Interactions with online perpetrators*

Almost two in three (64.6%) adolescents who experienced sexual extortion said it was perpetrated by someone they had met online and had never met in person (Figure 7). Of these adolescents, 90.9 percent said the perpetrator had contacted them first, and just 3.5 percent said they had made first contact with the perpetrator.

**Figure 7: Method of initial contact with perpetrator among adolescents who were sexually extorted (%) (*n*=220)**

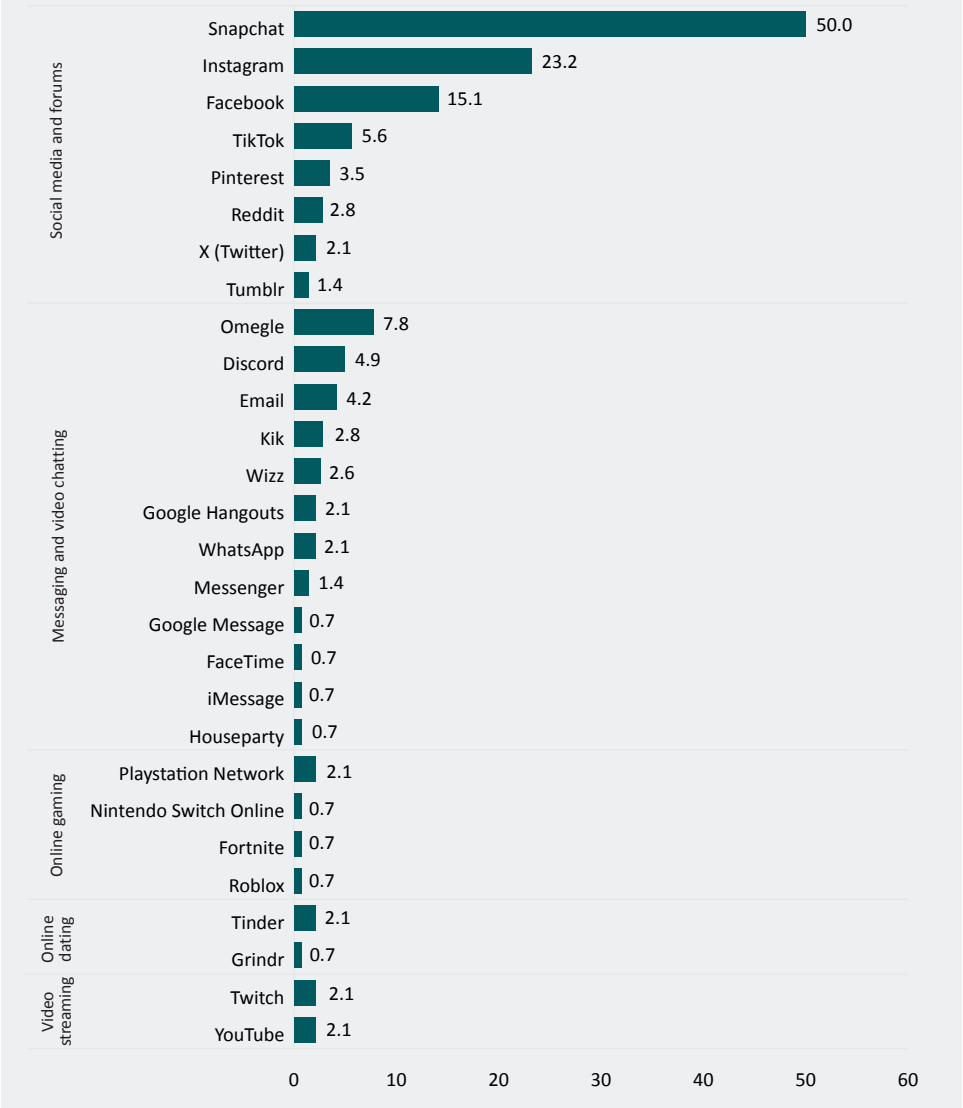



Note: Denominator includes three respondents (1.4%) who did not disclose if they had met a perpetrator online and eight respondents (5.6%) who chose not to disclose who made first contact

Source: AIC survey of adolescents, 2024 [computer file]

Among respondents who met the perpetrator online (*n*=142), half said they met on Snapchat (50.0%), followed by Instagram (23.2%), Facebook (14.1%), Omegle (7.8%), TikTok (5.6%) and Discord (4.9%); Figure 8). Social media platforms appeared to be the most common place of meeting an online perpetrator of sexual extortion.

**Figure 8: Online platforms where adolescents first had contact with perpetrators of sexual extortion (%) (*n*=142)**

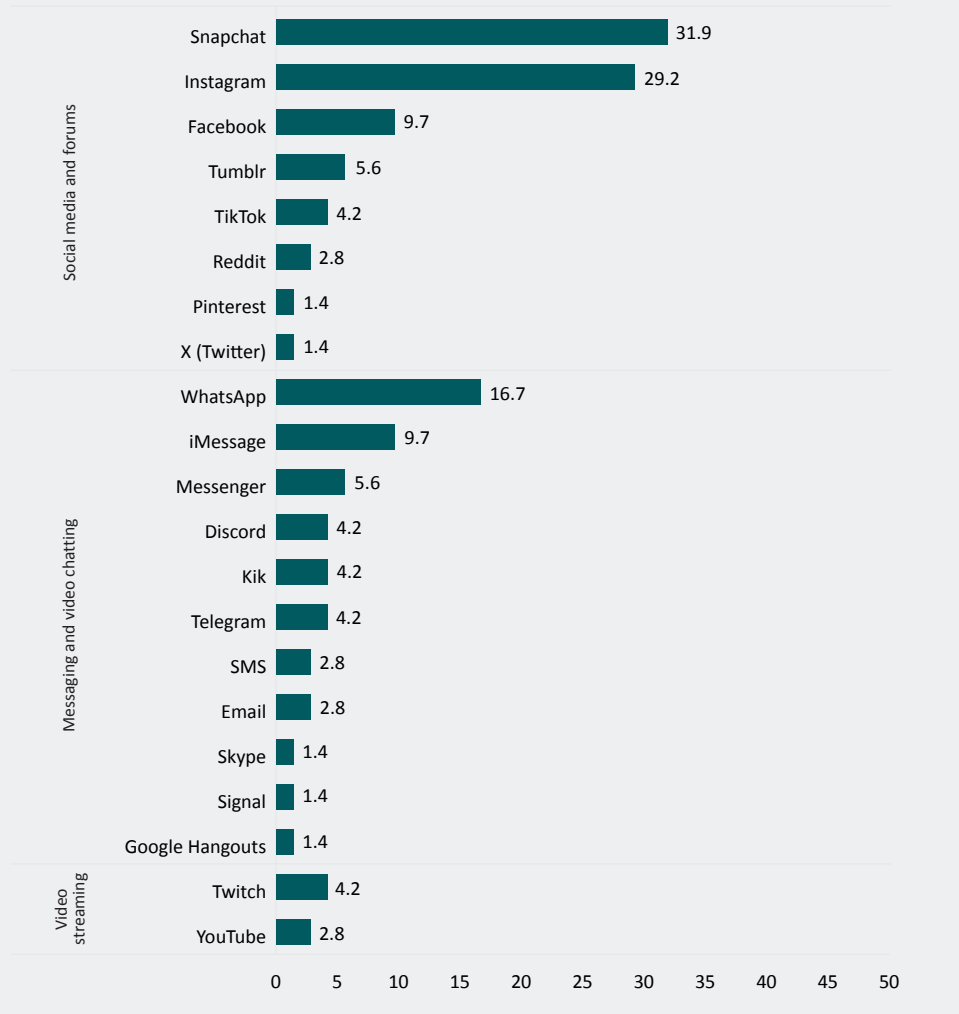



Note: Respondents could select multiple platforms, and four chose not to disclose. Platforms not selected by any respondent were not included in this figure—specifically, Signal, SMS, Telegram, Bumble, Hinge, OnlyFans, Steam, Call of Duty, Grand Theft Auto, Minecraft, Among Us, Xbox Live and Rocket League

Source: AIC survey of adolescents, 2024 [computer file]

Of the 142 respondents who had met the perpetrator online, half (50.7%) were asked by the perpetrator to switch to a new platform to continue communication. Typically, they were asked to switch to Snapchat (31.9%), Instagram (29.2%), WhatsApp (16.7%), Facebook (9.7%) or iMessage (9.7%; Figure 9).

**Figure 9: Online platforms adolescents were asked to switch to in order to continue communication with perpetrator of sexual extortion (%) (*n*=72)**

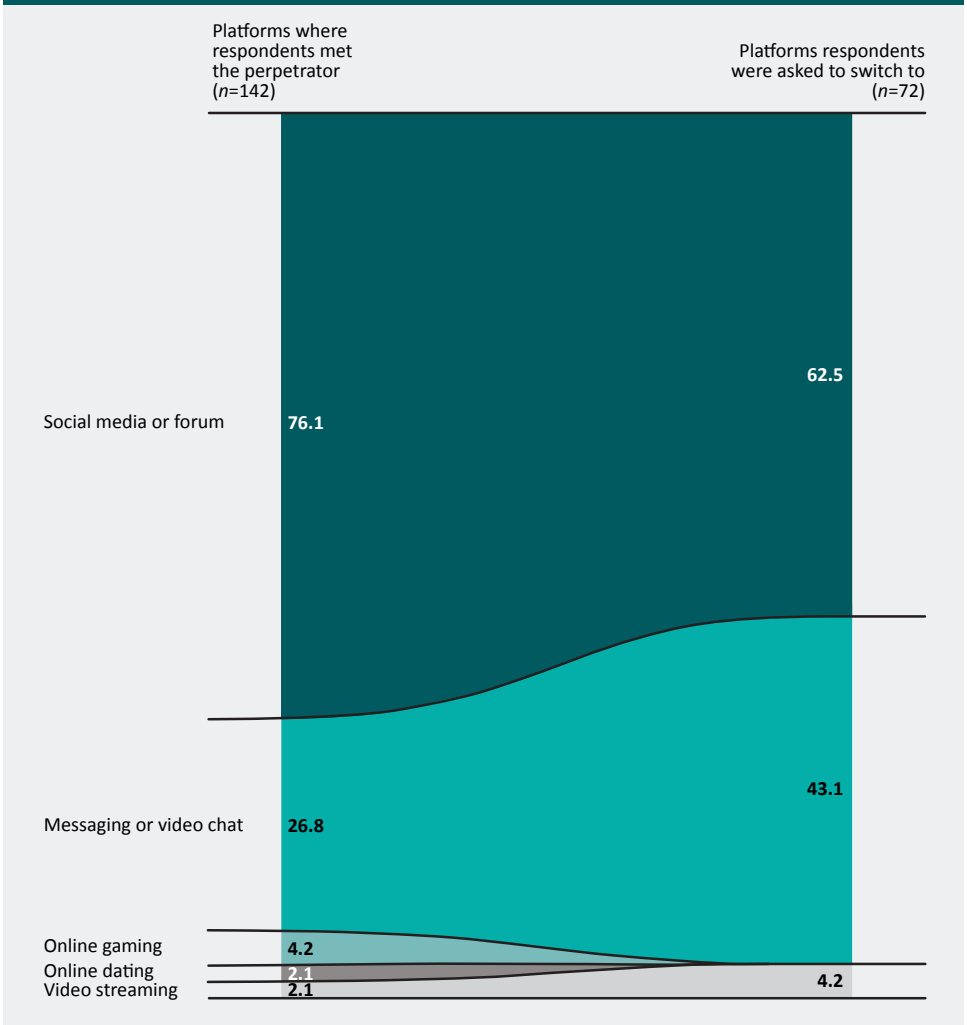



Note: Respondents could select multiple platforms, and one chose not to disclose. Platforms not selected by any respondent were not included in this figure

Source: AIC survey of adolescents, 2024 [computer file]

Respondents who were sexually extorted most commonly met their perpetrators on social media platforms (76.1%) and less commonly were asked by perpetrators to switch to social media platforms (62.5%; Figure 10). Conversely, respondents were more commonly asked to switch to messaging and video chatting platforms versus meeting perpetrators for the first time on these platforms (43.1% vs 26.8%). A minority of respondents first met the perpetrator on an online gaming platform (4.2%), but no respondents were asked to switch to these platforms. Finally, a small proportion of respondents were asked to switch to video streaming platforms (ie Twitch or YouTube; 4.2%), while even fewer first met their perpetrators on these platforms (2.1%).

**Figure 10: Difference in proportion of platforms used to initiate contact versus switched to for continued communication (%)**



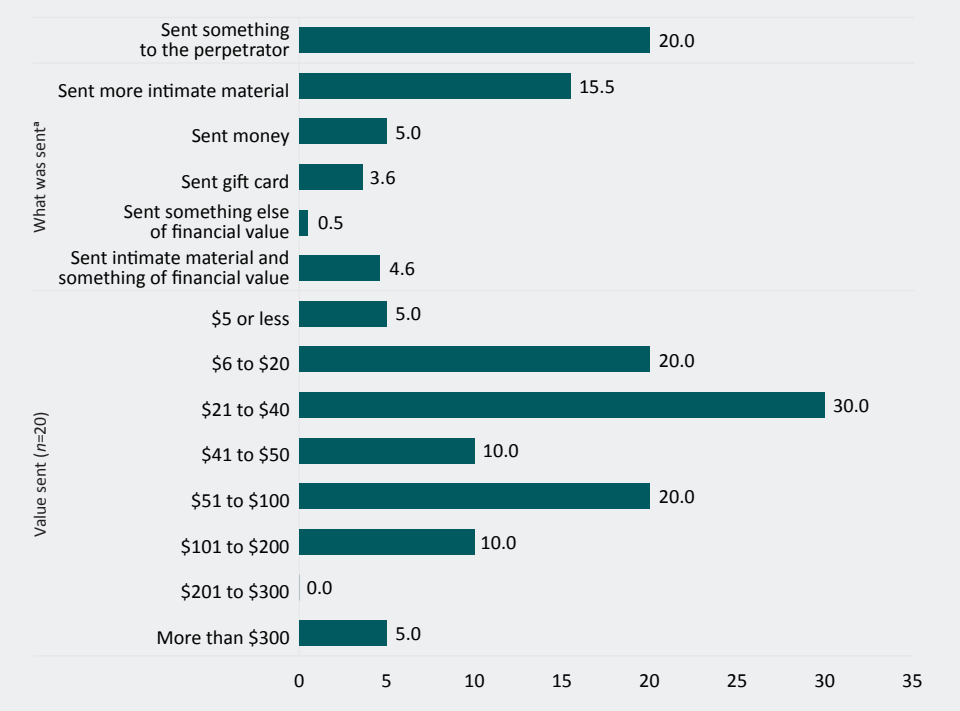

Note: Respondents could select multiple options meaning total does not equal 100 percent

Source: AIC survey of adolescents, 2024 [computer file]

### *Compliance with demands*

One in five adolescents who had been sexually extorted complied with the perpetrator's demands (20.0%; Figure 11), meaning 2.3 percent (44/1,953) of the adolescents in our sample experienced sexual extortion and complied with demands. Although roughly equal proportions of respondents received demands for money and intimate material, those who were sexually extorted were more likely to send intimate material (15.5%) than money (5.0%), while 3.6 percent sent a gift card and 0.5 percent sent something else of financial value (ie bank details or a book). Among those who sent money, a gift card or something else of financial value (*n*=20; 9.1% of those sexually extorted), most sent something worth $50 or less (*n*=13, 65.0%). One respondent sent something worth over $200.

**Figure 11: Prevalence and nature of compliance with sexual extortion demands (%) (*n*=220)**




a: Respondents could select multiple options for this survey question
Note: Denominators include respondents who did not disclose information
Source: AIC survey of adolescents, 2024 [computer file]

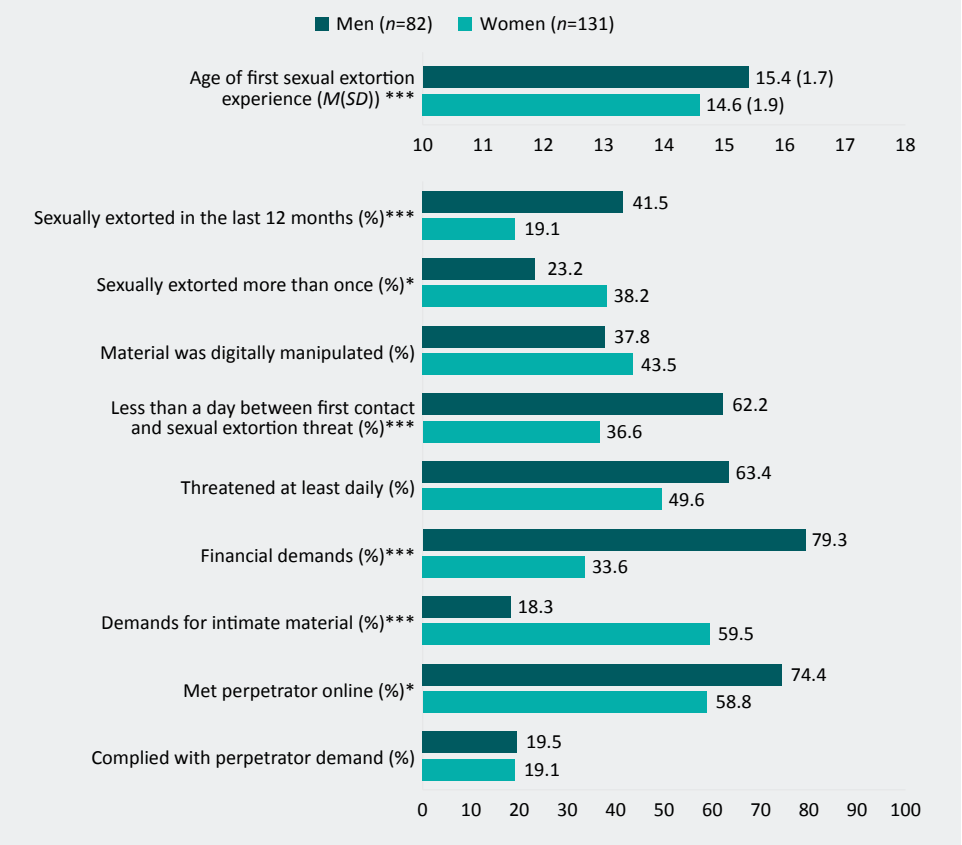
## Differences in experiences of sexual extortion

### *Gender differences*

We considered whether there were differences in how sexual extortion was experienced between genders. The average age of first sexual extortion was significantly younger for women (*M*=14.6, *SD*=1.9) than men (*M*=15.4, *SD*=1.7; $t(204)=3.0$, $p<0.01$; Figure 12). While women more commonly experienced sexual extortion in their lifetime (12.1% vs 10.1%; Figure 1), and were more likely to be revictimised (38.2% vs 23.2%; $\chi^2(1)=6.7$, $p<0.05$), men had more commonly experienced sexual extortion in the 12 months prior to the survey (41.5% vs 19.1%; $\chi^2(1)=13.3$, $p<0.001$). Further, men were significantly more likely to receive a sexual extortion threat within one day of first contact with a perpetrator (62.2% vs 36.6%; $\chi^2(1)=13.7$, $p<0.001$). They were also significantly more likely to be sexually extorted by someone who they had never met in person (74.4% vs 58.8%; $\chi^2(1)=4.8$, $p<0.05$).

Men were significantly more likely to receive financial demands from perpetrators (79.3% vs 33.6%; $\chi^2(1)=46.0$, $p<0.001$), while women were more likely to receive demands for additional intimate material (59.5% vs 18.3%; $\chi^2(1)=34.9$, $p<0.001$). Finally, there were no significant differences observed between genders on experiences of sexual extortion using digitally manipulated material, receiving threats at least daily, or compliance with demands. We were unable to include trans and gender diverse respondents in this analysis due to the small number that experienced sexual extortion ($n$=6).

**Figure 12: Differences in sexual extortion experiences by gender**




***statistically significant at $p<0.001$, *statistically significant at $p<0.05$

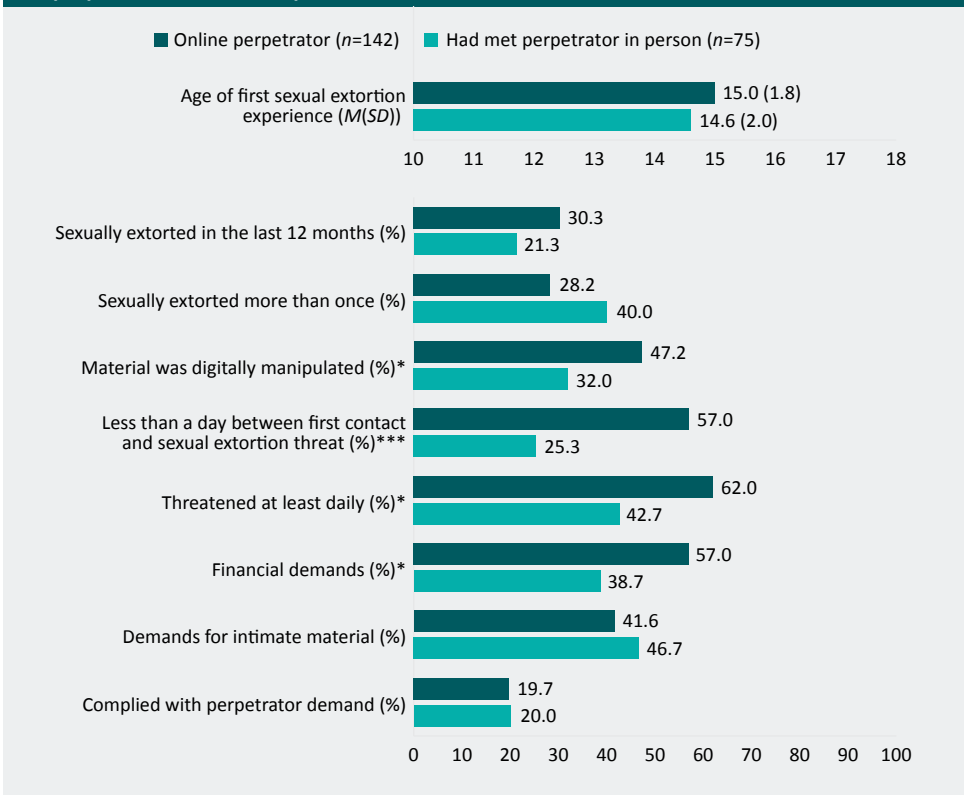
Note: Denominators include respondents who did not disclose information. Missing data were excluded for bivariate analyses. *M*=mean; *SD*=standard deviation

Source: AIC survey of adolescents, 2024 [computer file]

### *Online only sexual extortion versus known perpetrators*

Of those who had experienced sexual extortion, 64.6 percent said the perpetrator was a person they had only met online and never in person (Figure 13). Adolescents who had only met the perpetrator online were significantly more likely than other victims to have been extorted using digitally manipulated material (47.2% vs 32.0%; $\chi^2(1)=4.9$, $p<0.05$), to have less than a day between first contact and sexual extortion threat (57.0% vs 25.3%; $\chi^2(1)=18.6$, $p<0.001$), to have been threatened at least daily (62.0% vs 42.7%; $\chi^2(1)=6.4$, $p<0.05$), and to have received a financial demand (57.0% vs 38.7%; $\chi^2(1)=6.3$, $p<0.05$). Finally, there were no significant differences observed between the average age of respondents when first sexually extorted, revictimisation, recency of sexual extortion, demands for intimate material or the prevalence of complying with demands, based on how the victim knew the perpetrator.

**Figure 13: Differences in sexual extortion experiences based on whether adolescents had met the perpetrator online or in person**




***statistically significant at $p<0.001$, *statistically significant at $p<0.05$

Note: Denominators include respondents who were unsure or did not disclose information. These individuals were removed for bivariate analyses

Source: AIC survey of adolescents, 2024 [computer file]

## Discussion

Over one in 10 adolescents (11.3%) in our sample had experienced sexual extortion in their lifetime. The prevalence of sexual extortion in our study was higher than those from previous research using survey data from young people (eg 3.5 to 5%; Finkelhor, Turner & Colburn 2022; Patchin & Hinduja 2020), which may be explained by the use of differing definitions of sexual extortion, sample characteristics or recruitment methods. Despite these differences, the higher prevalence may indicate a substantial increase in sexual extortion among young people in recent years. Importantly, the present study used a community sample, and while these estimates cannot be generalised to other age groups, they are likely similar to the lifetime prevalence estimates of sexual extortion among adolescents across Australia between 16 and 18 years of age.

Similar to broader research on online child sexual exploitation (eg Teunissen et al. 2024), certain characteristics and online behaviours appeared to signal greater vulnerability to sexual extortion among adolescents. In particular, those who identified as sexually diverse and those who had a long-term health condition had higher rates of victimisation. Vulnerability was also heightened among those who had shared intimate material with someone, or had an online friend they had never met in person, an online dating profile, a larger number of social media profiles, and social media profiles set to public rather than private. Future research should continue to explore online activities and vulnerability to sexual extortion, especially given perpetrators infrequently used dating platforms to initiate contact. Increased visibility online, by virtue of using these platforms, may allow perpetrators to gather useful information before approaching adolescents on social media.

Importantly, sexual extortion appeared to vary substantially between cases, and in particular between genders, emphasising the heterogeneous nature of incidents. See Box 1 for a summary of sexual extortion experiences.

| Box 1: Summary of experiences among adolescents who had been sexually extorted (*n*=220) | |
|---|---|
| **27.3%** | Were sexually extorted in the last 12 months |
| **32.3%** | Were sexually extorted more than once |
| **57.7%** | Were sexually extorted when under the age of 16 |
| **41.4%** | Were sexually extorted with digitally manipulated materials |
| **45.9%** | Received a sexual extortion threat less than one day after first contact with the perpetrator |
| **55.0%** | Received sexual extortion threats daily |
| **50.5%** | Received a demand for something of financial value |
| **43.6%** | Received a demand for intimate material |
| **64.6%** | Were sexually extorted by someone they had met online and never in person |
| **20.0%** | Complied with sexual extortion demands |

Sexual extortion experiences differed based on whether adolescents had ever met the perpetrator in person, and on gender. Adolescents who had met a perpetrator solely online often said the extortion began soon after first contact, threats were frequent, and they had received financial demands. Our findings are consistent with evidence of an increase in sexual extortion perpetrated by strangers as a type of profit-motivated cybercrime, perhaps carried out by organised crime groups (Raffile et al. 2024; Thorn & NCMEC 2024).

Men more often said they had been sexually extorted recently, that the extortion began soon after meeting the perpetrator, that the extortion was financially motivated and that they had only met the perpetrator online. In contrast, women tended to be younger when first extorted, to experience greater time between meeting the perpetrator and the first threat, and to receive demands for intimate material. They also knew the perpetrator in person more often and were more frequently revictimised. Although it was beyond the scope of the survey to determine, these findings could indicate that young men are more likely to be targeted by organised crime groups seeking financial benefit (Raffile et al. 2024) and young women are more likely to be targeted by sexually motivated offenders. However, it is important to note that while men appeared to experience more financially motivated extortion, it is possible that intimate materials were extorted from women to be sold or distributed for financial benefit.

The prevalence of being sexually extorted by someone met solely online is much higher in our study than prior survey research (64.6% vs 13.6% Finkelhor, Turner & Colburn 2022; vs 27.2% Patchin and Hinduja 2020). Patchin and Hinduja (2020) found 24.2 percent of male victims and 17.4 percent of female victims in their sample had been sexually extorted by someone they only knew online, compared to 74.4 percent of male victims and 58.8 percent of female victims in our sample. While both studies reported a higher prevalence among men, the present study reported a higher overall prevalence among both genders, when compared with Patchin and Hinduja (2020). Although this difference may be partially attributed to the samples differing geographically (United States vs Australia) and in the age of respondents (12–17 years vs 16–18 years), it could also indicate a recent increase in sexual extortion perpetrated by strangers online.

## Implications

The prevalence of sexual extortion identified in the current study is concerning given the significant and varied impacts associated with victimisation (Wolbers et al. 2025). Our findings emphasise the importance of diverse and holistic prevention approaches to address the differing forms of sexual extortion experienced by adolescents. We identified several online platforms on which adolescents reported contact by perpetrators. While perpetrators commonly used popular social media platforms including Snapchat and Instagram to first contact adolescents, we found adolescents were being asked to switch to secondary platforms, including messaging platforms with end-to-end encryption (eg WhatsApp and iMessage). The use of encrypted communication to carry out sexual extortion introduces challenges for detection and deterrence. Yet, for most victims, sexual extortion took place on open and popular social media and messaging platforms—presenting opportunities for these platforms to take proactive action.

In particular, it would be useful to implement Safety by Design principles, which put user safety and rights at the centre of the design and development of online products and services (eSafety Commissioner 2024). Risk indicators such as the rapid creation of multiple accounts from the same device, attempted contact with many young people at once or the use of certain phrases or usernames associated with extortion should alert online platforms to suspicious accounts.

Separately, approaches that reduce the risk of adolescents being contacted by strangers online may have a significant impact on the occurrence of sexual extortion. Two in five adolescents who were sexually extorted reported that their intimate images had been digitally manipulated. This points to a need to ensure privacy settings for young users are set to their maximum by default, preventing malicious actors from contacting young people online or accessing their social media profiles. It also indicates a need to educate young people about the risks of sexual extortion and manipulation of images. In cases where the perpetrator is not an online stranger but a trusted friend or a current or former intimate partner, education approaches could equip adolescents with knowledge of risky online behaviours and where to seek help, as well as drawing attention to early intervention programs.

In addition to prevention, the scale of sexual extortion suggests victims may require support, including emotional or financial support and advice about what to do and how to remove intimate material posted online if the perpetrator follows through on their threats. Prior research has shown that sexual extortion can be a prolonged experience without intervention (Açar 2016; Alsoubai et al. 2022; Wolak et al. 2018), contributing to significant and ongoing negative impacts among victims. Similarly, reporting mechanisms must be accessible and age appropriate for adolescents, as research shows many young people do not report sexual extortion (Patchin & Hinduja 2020; Wolak et al. 2018). Adolescents may also require a reporting mechanism they recognise as safe, particularly where impacts may include legal repercussions (eg for producing and sharing CSAM; Wolbers et al. 2025). Given the present study examined bivariate comparisons between adolescents who had and had not been sexually extorted, a future paper will examine the factors that increase the risk of sexual extortion among adolescents in more depth and discuss the findings in greater detail.

## Conclusion

Using a national community-level sample, this research found that more than one in 10 adolescents residing in Australia may have experienced sexual extortion. These experiences appear to be highly diverse. There were considerable differences in victimisation based on the gender of adolescents and whether the perpetrator was someone met solely online or whether the victim had known them in person. Understanding differences in sexual extortion dynamics according to perpetrator behaviours and victim characteristics can help guide prevention and intervention activities. These approaches must be multifaceted and collaborative, bringing together law enforcement, government, financial institutions and, in particular, the tech sector to increase detection and intervention and reduce the opportunity for perpetration.

# References

*URLs correct as at November 2024*

Açar KV 2016. Sexual extortion of children in cyberspace. *International Journal of Cyber Criminology* 10(2): 110–126. https://doi.org/10.5281/zenodo.163398

ACCCE—see Australian Centre to Counter Child Exploitation

Alsoubai A, Song J, Razi A, Naher N, De Choudhury M & Wisniewski PJ 2022. From 'friends with benefits' to 'sextortion:' A nuanced investigation of adolescents' online sexual risk experiences. *Proceedings of the ACM on Human-Computer Interaction* 6(CSCW2): article 411. https://doi.org/10.1145/3555136

Australian Bureau of Statistics (ABS) 2023. National, state and territory population, September 2022. https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/sep-2022

Australian Centre to Counter Child Exploitation 2022. Sextortion: Messaging kit. https://www.accce.gov.au/sites/default/files/2023-11/Sextortion_messaging_kit_October2023.pdf

Branje S, De Moor EL, Spitzer J & Becht AI 2021. Dynamics of identity development in adolescence: A decade in review. *Journal of Research on Adolescence* 31(4): 908–927. https://doi.org/10.1111/jora.12678

C3P—see Canadian Centre for Child Protection

Canadian Centre for Child Protection 2024. Online harms: Sextortion. https://www.cybertip.ca/en/online-harms/sextortion/

Crone EA & Konijn EA 2018. Media use and brain development during adolescence. *Nature Communications* 9(1): 588. https://doi.org/10.1038/s41467-018-03126-x

Cross C, Holt K & O'Malley RL 2022. "If u don't pay they will share the pics": Exploring sextortion in the context of romance fraud. *Victims & Offenders* 18(7): 1194–1215. https://doi.org/10.1080/15564886.2022.2075064

Duell N et al. 2018. Age patterns in risk taking across the world. *Journal of Youth and Adolescence* 47: 1052–1072. https://doi.org/10.1007/s10964-017-0752-y

eSafety Commissioner 2024. Safety by Design. https://www.esafety.gov.au/industry/safety-by-design

eSafety Commissioner 2023a. *Sexual extortion and child abuse reports almost triple*. Media release, 25 May. https://www.esafety.gov.au/newsroom/media-releases/sexual-extortion-and-child-abuse-reports-almost-triple

eSafety Commissioner 2023b. *1 in 8 children coerced into producing child sexual abuse material remotely*. Media release, 5 September. https://www.esafety.gov.au/newsroom/media-releases/1-in-8-children-coerced-into-producing-child-sexual-abuse-material-remotely

Finkelhor D, Turner H & Colburn D 2023. Which dynamics make online child sexual abuse and cyberstalking more emotionally impactful: Perpetrator identity and images? *Child Abuse & Neglect* 137: 106020. https://doi.org/10.1016/j.chiabu.2023.106020

Finkelhor D, Turner H & Colburn D 2022. Prevalence of online sexual offenses against children in the US. *JAMA Network Open* 5(10): e2234471. https://doi.org/10.1001/jamanetworkopen.2020.34471

Gewirtz-Meydan A, Walsh W, Wolak J & Finkelhor D 2018. The complex experience of child pornography survivors. *Child Abuse & Neglect* 80: 238–248. https://doi.org/10.1016/j.chiabu.2018.03.031

Internet Watch Foundation 2023. *How AI is being abused to create child sexual abuse imagery*. Internet Watch Foundation. https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/

Liggett R 2019. Exploring online sextortion. *Family & Intimate Partner Violence Quarterly* 11(4): 45–56. https://www.civicresearchinstitute.com/online/article_abstract.php?pid=6&iid=1395&aid=9106

Madigan S, Ly A, Rash CL, Van Ouytsel J & Temple JR 2018. Prevalence of multiple forms of sexting behavior among youth: A systematic review and meta-analysis. *JAMA Pediatrics 172*(4): 327–335. https://doi.org/10.1001/jamapediatrics.2017.5314

National Center for Missing and Exploited Children 2024. NCMEC releases new sextortion data. https://www.missingkids.org/blog/2024/ncmec-releases-new-sextortion-data

National Center for Missing and Exploited Children 2016. *Trends identified in CyberTipline sextortion reports*. National Center for Missing & Exploited Children. https://www.missingkids.org/content/dam/missingkids/pdfs/ncmec-analysis/sextortionfactsheet.pdf

O'Malley RL & Holt KM 2022. Cyber sextortion: An exploratory analysis of different perpetrators engaging in a similar crime. *Journal of Interpersonal Violence* 37(1–2): 258–283. https://doi.org/10.1177/0886260520909186

Patchin JW & Hinduja S 2024. The nature and extent of youth sextortion: Legal implications and directions for future research. *Behavioral Sciences & the Law* 42(4): 401–416. https://doi.org/10.1002/bsl.2667

Patchin JW & Hinduja S 2020. Sextortion among adolescents: Results from a national survey of U.S. youth. *Sexual Abuse* 32(1): 30–54. https://doi.org/10.1177/1079063218800469

Raffile P, Goldenberg A, McCann C & Finkelstein J 2024. *A digital pandemic: Uncovering the role of 'Yahoo Boys' in the surge of social media-enabled financial sextortion targeting minors*. Network Contagion Research Institute. https://networkcontagion.us/reports/yahoo-boys/

Teunissen C, Thomsen D, Napier S & Boxall H 2024. Risk factors for receiving requests to facilitate child sexual exploitation and abuse on dating apps and websites. *Trends & issues in crime and criminal justice* no. 686. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77291

Thorn & National Centre for Missing and Exploited Children (NCMEC) 2024. *Trends in financial sextortion: An investigation of sextortion reports in NCMEC cybertipline data*. https://www.thorn.org/research/library/financial-sextortion/

Wolak J, Finkelhor D, Walsh W & Treitman L 2018. Sextortion of minors: Characteristics and dynamics. *Journal of Adolescent Health* 62(1): 72–79. https://doi.org/10.1016/j.jadohealth.2017.08.014

Wolbers H, Cubitt T, Carter R & Napier S 2025. The impacts of sexual extortion on minors: A systematic review. *Trends & issues in crime and criminal justice* no. 710. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77789

This work was originally published as *Trends & issues in crime and criminal justice* no. 712. It is available online at **https://doi.org/10.52922/ti77819**.

# 11. The impacts of sexual extortion on minors: A systematic review

Heather Wolbers, Timothy Cubitt, Renee Carter and Sarah Napier

The rapid growth and development of online spaces has resulted in the emergence of new forms of technology-facilitated sexual violence (Cross, Holt & O'Malley 2022). One of these is sexual extortion—a form of blackmail in which a perpetrator threatens to distribute sexual images or videos of a victim unless they comply with certain demands (Açar 2016; Alsoubai et al. 2022; Chiang & Grant 2019; Henry, Flynn & Powell 2019; Liggett 2019). Demands are often for money, additional intimate materials, or sexual contact with the victim (Eaton, Ramjee & Saunders 2023; Kopecký 2017; Patchin & Hinduja 2020).

Although sexual extortion is perpetrated against both adults and children, children appear particularly vulnerable to this crime type (O'Malley & Holt 2022; Wittes et al. 2016; Wolak & Finkelhor 2016; Wolak et al. 2018). This is thought to be in part due to their increased time spent online, coupled with their elevated risk tolerance and lower social and developmental maturity (Açar 2016; Alsoubai et al. 2022; Douglass et al. 2018; Patchin & Hinduja 2020).

Sexual extortion is a significant online threat to children (Europol 2015). Similar to many forms of sexual violence and technology-enabled offences, under-reporting is common due to embarrassment, fear of retribution from the offender, and concern about punishment for behaving sexually online (Alsoubai et al. 2022; Wolak & Finkelhor 2016). The Australian Centre to Counter Child Exploitation (2023) receives hundreds of reports each month of financially motivated sexual extortion of minors in Australia. However, research from the United States suggests that many child and young adult victims of sexual extortion do not report their experiences (Patchin & Hinjuda 2020; Wolak et al. 2018), meaning the true rate of victimisation is expected to be much higher. For example, a representative survey of United States school students showed that five percent of 12- to 17-year-olds had been sexually extorted; of these, only 6.9 percent said they reported it to police (Patchin & Hinduja 2020).

The sexual extortion of children contributes to the availability of child sexual abuse material (CSAM) online. Instances of sexual extortion involve the production and distribution of CSAM, as offenders groom and coerce children to self-generate sexual photos and videos. CSAM victims can experience helplessness, shame, anxiety, grief, betrayal and revictimisation through the production of their abusive content and its distribution online (Canadian Centre for Child Protection 2017; Salter et al. 2021). Like other online crimes against children, sexual extortion of minors is new and emerging, and thus little research has examined the impacts of this crime on children.

## Objectives of this review

To ensure that the support needs of minors are met following experiences of victimisation and to understand the seriousness of this crime from a policy perspective, an accurate understanding of the impacts of sexual extortion on both their short- and long-term wellbeing must be developed. To do this, we use a systematic review methodology to assess what is known about the impacts of sexual extortion on individuals under the age of 18. Systematic reviews use search, screening and analytic techniques to collate evidence about a topic. Importantly, within the past decade, sexual extortion has been the focus of research and has been acknowledged as a significant and widespread problem among police, government and the not-for-profit sectors.

While sexual extortion overlaps with other forms of child sexual exploitation and cybercrime, it is currently unclear whether victims of sexual extortion experience different, or additional, impacts to their wellbeing. In particular, certain aspects of this crime make it unique, such as the use of coercion, threats and blackmail. Intimate materials can be obtained in a variety of ways (eg grooming or coercion, taking photos without consent, hacking), and threats of dissemination are made for a variety of motives (eg financial gain or sexual gratification). For these reasons, prior research into the impacts of other forms of child sexual abuse may not capture the totality of impacts experienced by victims of sexual extortion. Given the increasing use of the internet by minors (Anderson, Faverio & Gottfried 2023), and their vulnerability to technology-facilitated offending (Teunissen et al. 2024), understanding the impacts of sexual extortion is crucial in guiding interventions that protect children from online harms and informing future research into this growing crime type.

To assist in a broader study involving several separate systematic reviews on topics related to sexual extortion of minors, a pilot literature search was conducted in July 2023 by the Australian Institute of Criminology's JV Barry Library. This search found that the empirical research on prevalence and offender methods and tactics was not up to date with the rapidly changing trends in sexual extortion against minors demonstrated in police and not-for-profit reports (eg Australian Centre to Counter Child Exploitation 2022; National Center for Missing and Exploited Children (NCMEC) 2023). We therefore focused this systematic review on the *impacts* of sexual extortion on minors, a topic that is less likely to be influenced by rapid changes in offending methodologies. Our selection criteria captured academic and grey literature that included primary data on sexual extortion perpetrated against minors and the impacts of victimisation.

# Methods

## Systematic search strategy

This study draws on all available sources examining the impacts of sexual extortion where the victim was under 18 years of age, including impacts on mental health, social activities, finances, and future victimisation or perpetration. A full systematic literature search was conducted in December 2023. The search was limited to studies written in English that were published in the 20 years prior to the search (2003 onwards). Four databases were searched: the JV Barry Library catalogue, ProQuest, DeepDyve and EBSCO. The websites of relevant organisations such as the International Centre for Missing and Exploited Children, NCMEC and the Office of the eSafety Commissioner were also searched. We undertook additional screening of the first 200 results from Google Scholar and 100 links from Google to identify any sources that may not have appeared in the databases and websites. Finally, the reference lists of records that were ultimately included in the review were checked for any additional sources (ie backchaining).

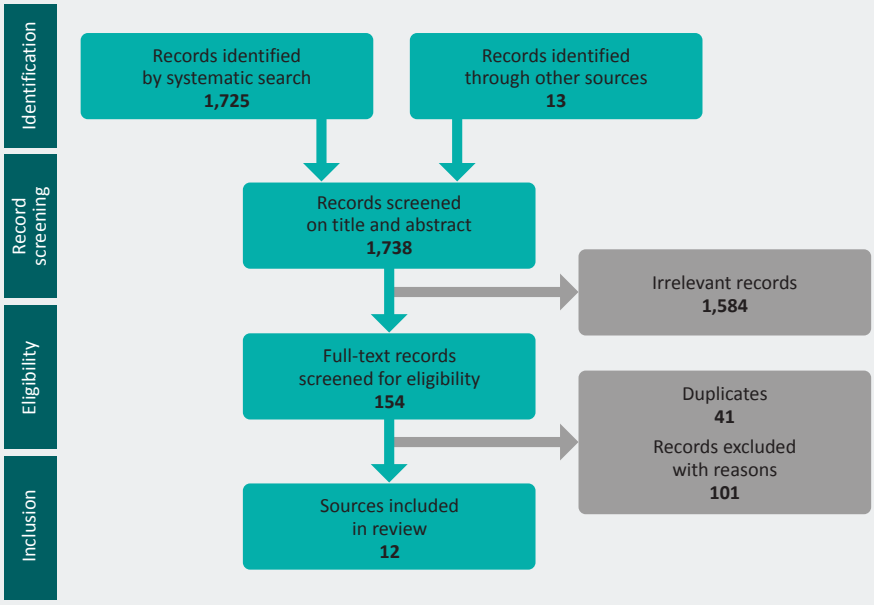
Assisted by the initial pilot search, we developed a list of search terms to guide this systematic literature search. The search terms used combined keywords from three categories:

- *sexual extortion* (sextort* OR “sexual extort*” OR “cyber extort*” OR “online extort*” OR “cyber blackmail” OR “online blackmail”);
- *age of the victim* (child* OR “young person” OR teenage* OR adolescen* OR youth OR minor OR “young adult”); and
- *impacts* (impact OR harm OR trauma OR mental health OR psychological outcome OR outcome OR suicid* OR consequence OR wellbeing OR financ* OR cost OR loss OR education OR employment OR long term OR short term OR ramification OR revictimi?ation OR “repeat victim*” OR polyvictimi?ation OR “victim-offender overlap” OR “victim to offender cycle” OR “criminal justice involvement”).

## Screening process

The systematic search yielded 1,725 studies. An additional 13 studies were identified through wider reading, resulting in identification of an initial 1,738 records. Titles and abstracts were screened to exclude irrelevant studies (*n*=1,584). This screening yielded 154 studies, of which 41 were identified as duplicates and subsequently removed. The remaining 113 sources were assessed for eligibility against the selection criteria with full-text screening, and 101 were excluded because they did not analyse primary data (eg literature reviews), examine impacts on child victims of sexual extortion or provide sufficient detail on the data source. In total, the systematic search of literature yielded 12 sources providing primary information on the impact of sexual extortion on victims who were minors (see Figure 1).

**Figure 1: Literature screening process**




## Limitations

The sample of 12 studies contained a variety of measures of sexual extortion and its subsequent impacts, and featured a variety of data sources and sampling methodologies. Not all studies that met the inclusion criteria reported sufficient quantitative data to calculate standardised effect sizes, while some used qualitative research methods. Ultimately, this meant that conducting a meta-analysis of these studies was not possible (Moher et al. 2009). Studies also varied by cultural context, meaning the interpretation of findings contained culture-specific nuance. Accordingly, the following sections provide a qualitative synthesis of the eligible studies with discussion of the strength or magnitude of the effect, where the sample size allowed. These limitations mean caution should be employed when comparing the results between studies.

# Results

## Characteristics of eligible studies

Studies that met the inclusion criteria were primarily from the United States (*n*=8), but were also drawn from Israel (*n*=2), Australia (*n*=1), and the United Kingdom (*n*=1; See Table 1). Four studies used qualitative methods, three used quantitative methods and five used both. Despite searching for literature over the last 20 years, published research was only available from 2016 onward, highlighting that sexual extortion of minors is a current and emerging issue.

**Table 1: Studies included in the systematic review**

| Study | Country | Data source | Method |
|---|---|---|---|
| **Peer-reviewed literature** | | | |
| Alsoubai et al. (2022) | United States[a] | 45,955 posts (3,767 for harassment/ sextortion) made by 13- to 17-year-olds on an online peer support platform | Mixed |
| Dolev-Cohen (2023) | Israel | Semi-structured in-depth interviews with 32 Arab educational counsellors, 8 with instances of sexual extortion | Qualitative |
| Dolev-Cohen, Nezer & Zumt (2023) | Israel | Semi-structured in-depth interviews with 20 Jewish or Arab school counsellors who treated female students who were sexually extorted online | Qualitative |
| Finkelhor, Turner & Colburn (2023) | United States | Nationally representative online survey of 2,639 young adults asked retrospectively about childhood; 81 were victims of sexual extortion | Quantitative |
| Nilsson et al. (2019) | United Kingdom[a] | Publicly available online material on 3 cases of teenagers who committed suicide after being sexually extorted | Qualitative |
| O'Malley & Holt (2022) | United States | Open-source news articles and court documents regarding 152 cases of cyber sexual extortion which occurred from 2008 to 2018; 80 were minor focused offenders | Qualitative |
| Patchin & Hinduja (2020) | United States | Nationally representative survey of 5,568 middle and high school students administered from Aug to Oct 2016 | Quantitative |
| Wolak et al. (2018) | United States | Online survey of 1,628 people aged 18–25 years who had experienced sexual extortion; 572 recent cases were under 18 when threats began | Quantitative |
| **Grey literature** | | | |
| AUSTRAC (2022) | Australia | Financial transactions for sexual extortion | Mixed |
| NCMEC (2016) | United States | 1,428 reports of sexual extortion of a child received by CyberTipline from Oct 2013 to Apr 2016 | Mixed |
| Thorn (2019) | United States | Online survey of 2,097 people aged 13–25 years who had experienced sexual extortion | Mixed |
| Wolak & Finkelhor (2016) | United States | Online survey of 1,631 people aged 18–25 years who had experienced sexual extortion; 740 were a minor when threats began | Mixed |

a: Sample was multi-national

While there were marginal differences in how sexual extortion was described, all included studies featured the threat to disseminate intimate material in their definition (Table 2). Some studies identified the intention of the extortion activity, most commonly being to obtain money, additional intimate material or sexual contact (Alsoubai et al. 2022; Dolev-Cohen 2023; Finkelhor, Turner & Colburn 2023; NCMEC 2016; Patchin & Hinduja 2020; Wolak et al. 2018). However, intent was more broadly conceptualised in others (Dolev-Cohen, Nezer & Zumt 2023; Nilsson et al. 2019; O'Malley & Holt 2022; Patchin & Hinduja 2020; Thorn 2019; Wolak & Finkelhor 2016). Several studies focused on technology-facilitated sexual extortion (Alsoubai et al. 2022; Nilsson et al. 2019; O'Malley & Holt 2022), or financial sexual extortion (AUSTRAC 2022).

**Table 2: Conceptual definitions of sexual extortion used by included studies**

| Study | Definition |
|---|---|
| AUSTRAC (2022) | Coercion of Australian children into producing explicit images and extorting them for money |
| Alsoubai et al. (2022) | Threats to expose sexual images to coerce victims to provide additional pictures, sex, or other favours |
| Dolev-Cohen (2023) | Threat of dissemination of sexual images of the subject and extortion to force the victim to send additional sexual images, pay money, engage in sexual intercourse, or to prevent the victim from leaving a relationship |
| Dolev-Cohen, Nezer & Zumt (2023) | A person threatens another with the distribution of sexual content on the internet, to obtain more pictures or videos, money, or have some other demand met |
| Finkelhor, Turner & Colburn (2023) | A threat to disseminate sexual images in their possession to obtain money, additional pictures, or other sexual activities |
| NCMEC (2016) | Utilisation of non-physical forms of coercion, such as blackmail, to acquire sexual content of the child, obtain money from the child or engage in sex with the child |
| Nilsson et al. (2019) | Facilitated by the use of online environments where perpetrators gain the trust of vulnerable individuals in order to obtain sexually explicit material and then use it to coerce victims for the purposes of sexual, personal or financial gain |
| O'Malley & Holt (2022) | Occurring through a technological medium, threats to distribute intimate images or videos unless specific demands were met |
| Patchin & Hinduja (2020) | The threatened dissemination of explicit, intimate, or embarrassing images of a sexual nature without consent, usually for the purpose of procuring additional images, sexual acts, money, or something else |
| Thorn (2019) | The threat to expose a sexual image in order to make a person do something or for other reasons, such as revenge or humiliation |
| Wolak & Finkelhor (2016) | Threats to expose a sexual image in order to make a person do something or for other reasons, such as revenge or humiliation |
| Wolak et al. (2018) | Threats to expose sexual images to coerce victims to provide additional pictures, sex, or other favours |

## Impacts of sexual extortion victimisation on minors

The impacts of sexual extortion on minors were diverse, ranging from immediate effects through to those that lasted or manifested years later. These impacts could be separated into five categories: emotional and psychological wellbeing, socio-environmental, financial, ongoing and poly-victimisation, and offending or criminal justice involvement (Table 3).

**Table 3: Summary of the impacts of sexual extortion victimisation on minors**

| | |
|---|---|
| **Impacts on emotional and psychological wellbeing** | |
| Poor mental wellbeing (4, 7, 8) | Feelings led to non-reporting of extortion (10, 12) |
| Feelings of fear, anger, shame, helplessness, hopelessness or loneliness (2, 4, 5, 6, 7, 10) | Required care from a mental or medical health professional (10, 12) |
| Anxiety and depression (4, 5, 6) | Self-harm and suicide (6, 7) |
| **Socio-environmental impacts** | |
| Experienced social isolation (3, 5, 10, 12) | Had to change or take time off school (3, 10, 12) |
| Experienced reputational damage (3, 7) | Lost or had to change their job (12) |
| Experienced or risked damage to family relationships (3, 4) | Had to move to a different location (10, 12) |
| Had to modify their online presence (10) | |
| **Financial impacts** | |
| Incurred financial costs (12) | When there were demands for money, victims usually sent between $50 and $1,500 (1) |
| **Ongoing victimisation and poly-victimisation** | |
| Sexual extortion could be prolonged, with threats received for over a month (12) | Threats were carried out by perpetrators in around 1 in 4 cases (9, 12) |
| Complying with the demands could lead to continuation of sexual extortion (10, 12) | Sexual extortion was often accompanied by at least one other form of victimisation (9, 12) |
| Additional, different threats often accompanied extortion of intimate materials (12) | |
| **Offending and criminal justice involvement** | |
| Criminal proceedings considered against victims for CSAM offences (11) | Victims of sexual extortion had often been perpetrators of sexual extortion as well (9) |
| Victims were coerced by perpetrators into sexually offending (8) | |

Note: 1—AUSTRAC (2022); 2—Alsoubai et al. (2022); 3—Dolev-Cohen (2023); 4—Dolev-Cohen, Nezer & Zumt (2023); 5—Finkelhor, Turner & Colburn (2023); 6—NCMEC (2016); 7—Nilsson et al. (2019); 8—O'Malley & Holt (2022); 9—Patchin & Hinduja (2020); 10—Thorn 2019; 11—Wolak & Finkelhor (2016);
12—Wolak et al. (2018)

### *Impacts on emotional and psychological wellbeing*

Most of the examined studies included information on the emotional and psychological wellbeing of victims following sexual extortion. Poor general mental wellbeing and distress was reported in three studies (Dolev-Cohen, Nezer & Zumt 2023; Nilsson et al. 2019; O'Malley & Holt 2022). However, some studies detailed specific emotional and psychological impacts of sexual extortion on minors. Several studies found that minors who were sexually extorted experienced fear (Alsoubai et al. 2022; Finkelhor, Turner & Colburn 2023; NCMEC 2016; Nilsson et al. 2019; Thorn 2019), or shame and embarrassment (Alsoubai et al. 2022; Dolev-Cohen, Nezer & Zumt 2023; Finkelhor, Turner & Colburn 2023; Nilsson et al. 2019; Thorn 2019). Three studies found that sexual extortion was associated with feelings of helplessness or hopelessness (Alsoubai et al. 2022; NCMEC 2016; Nilsson et al. 2019), and anxiety or depression (Dolev-Cohen, Nezer & Zumt 2023; Finkelhor, Turner & Colburn 2023; NCMEC 2016). Two studies found that minors who were sexually extorted experienced loneliness and sadness (Dolev-Cohen, Nezer & Zumt 2023; Finkelhor, Turner & Colburn 2023), while one study found that victimisation was associated with anger (Finkelhor, Turner & Colburn 2023).

Nilsson et al. (2019) described how victims felt ongoing fear of the next demand from the perpetrator, and fear of the reputational damage associated with people in their life finding out about the sexual extortion. Some victims reported that law enforcement had limited ability to respond to these offences, or feeling that they may not be taken seriously, coupled with a fear that their content could be shared online, resulting in feelings of helplessness and hopelessness. They often blamed themselves for the victimisation, and when the perpetrator shared the content victims experienced a deep sense of humiliation and shame. Two studies described how these negative emotions, particularly fear and shame, were the reason victims chose not to disclose the sexual extortion experience to anyone (Thorn 2019; Wolak et al. 2018). In two studies that reported the relevant information, about a third of victims sought mental or medical health advice as a direct result of victimisation (1 in 3, Thorn 2019; 28.7%, Wolak et al. 2018).

Importantly, NCMEC (2016) reported that four percent of all sexual extortion victims who submitted reports had engaged in self-harm or threatened or attempted suicide after the victimisation. These impacts were further supported by Nilsson et al. (2019), who focused on three teenagers who lost their lives to suicide following sexual extortion. These findings illustrated that sexual extortion can have extreme consequences for individuals under the age of 18.

### *Socio-environmental impacts*

It was common for victims to report becoming socially isolated and losing friends due to sexual extortion (Dolev-Cohen, Nezer & Zumt 2023; Thorn 2019; Wolak et al. 2018). Wolak et al. (2018) showed that 46 percent of minors who were sexually extorted lost a relationship with a friend or family member, while Finkelhor, Turner and Colburn (2023) revealed that 96 percent of sexual extortion victims in their sample felt they were alone, all of whom felt like they could not trust people. Victims of sexual extortion also reported having to modify their online identity, specifically closing accounts (38%) and changing usernames (36%; Thorn 2019).

Many experienced what they perceived to be damage to their reputation (Dolev-Cohen, Nezer & Zumt 2023; Nilsson et al. 2019). Furthermore, interviews with school counsellors in Israel highlighted that minors who were sexually extorted had additional social and safety concerns about their families finding out about the sexual extortion (Dolev-Cohen 2023; Dolev-Cohen, Nezer & Zumt 2023). These findings demonstrated that the impacts of sexual extortion could be influenced by context and could be culture-specific.

The literature additionally showed that victims experienced disruptions to other aspects of their lives. Many missed school or had to change schools entirely (13.6%, Wolak et al. 2018; 1 in 8, Thorn 2019; see also Dolev-Cohen, Nezer & Zumt 2023). Similarly, in one study disruptions to employment were reported by around eight percent of victims (Wolak et al. 2018). Two studies found that, in one in 10 cases, victims had to leave their community or town by moving their residential location (Thorn 2019; Wolak et al. 2018). These findings suggest that the impacts of sexual extortion extended beyond emotional or psychological, to material changes in circumstances that could relate to friends, family, education, employment and housing.

### *Financial impacts*

While sexual extortion was not always financially motivated (Thorn 2019; Wolak et al. 2018), we identified two studies that discussed financial impacts on victims of sexual extortion. In the first study, four percent of minors who were sexually extorted received demands for money, and a further 7.9 percent incurred some kind of financial loss (Wolak et al. 2018). The second study, which focused solely on financial sexual extortion, found the average child or adolescent victim paid between $50 and $1,500, either as a lump sum or incrementally, before they ran out of funds (AUSTRAC 2022). However, some victims paid up to $10,000. While not all sexual extortion involved demands for money, among those that did the costs could be significant. Further, when the demands for money were not complied with, or the victim no longer had access to finances, victimisation still involved financial costs such as help-seeking and psychological care.

### *Ongoing victimisation and poly-victimisation*

Several studies discussed coercion and threats as a continuing, often drawn out impact of sexual extortion. For more than a quarter of victims (26.9%) in one study, threats continued for more than six months and just 17.0 percent said the threats lasted for less than a week (Wolak et al. 2018). In another study threats were commonly received daily (47%) and repeatedly, with about half of victims receiving between one and nine threats each day (52%; Thorn 2019). Importantly, complying with a perpetrator's demands did not often appear to stop victimisation. For almost two-thirds of victims who complied (64%), the threats continued (Thorn 2019).

Perpetrator threats went beyond sharing the extorted material with others and were often accompanied by threats of offline harm such as the victim getting in trouble at work or school, or that the perpetrator would harass or stalk the victim in person (Wolak et al. 2018). Some victims received threats that they would be assaulted, sexually assaulted, killed or otherwise hurt if they did not comply (29.2%). Two studies reported that in around a quarter of cases, threats were carried out and the victim's intimate materials were shared (Patchin & Hinduja 2020; Wolak et al. 2018). Wolak et al. (2018) further found that threats, regardless of the type, were carried out in almost half of the cases considered (47.2%).

During sexual extortion incidents, including when threats were being carried out or as part of intimate partner abuse, victims were harmed in a variety of ways. Victims experienced their accounts being hacked or fake accounts being made; being stalked or harassed online and in person; getting in trouble at school or work; having their family, friends, or pets harmed; and being beaten, raped or physically hurt (Patchin & Hinduja 2020; Wolak et al. 2018). Among the studies that measured ongoing and poly-victimisation, findings suggested that these impacts were common, and while they varied case by case they often involved serious and prolonged threats of physical or non-physical offending against the victim.

### *Offending and criminal justice involvement*

Three studies included information on offending or criminal justice involvement among minors who had been victims of sexual extortion. In the first, as minors were often tricked or coerced into producing and sharing intimate materials of themselves, some victims reported that police had considered taking action against them for CSAM offences (Wolak & Finkelhor 2016). These cases were in the minority and featured in the oldest study in the present review. Nevertheless, this is an important finding as it highlights that sexual extortion can position the victim as a perpetrator, creating barriers to reporting and help-seeking.

The second study found that during the sexual extortion incidents some victims were coerced into committing sexual offences. For example, in some cases the perpetrator demanded that the victim engage in the sexual abuse of other children, or engage in other sexually deviant behaviour (eg bestiality; O'Malley & Holt 2022). This not only put victims at risk of criminal charges but was also used by perpetrators to further intimidate and scare the victim with threats of serious legal action.

The third study indicated an overlap between victimisation and perpetration of sexual extortion. Patchin and Hinduja (2020) found that, among minors who perpetrated sexual extortion, 73 percent reported being victims of sexual extortion in the past. Further, 44 percent of victims of sexual extortion in the sample reported perpetrating sexual extortion. Without time-series data, it is difficult to determine the exact relationship between victimisation and perpetration of sexual extortion among minors in Patchin and Hinduja's study. However, their findings suggest a potential pathway from perpetration to victimisation and vice versa. The authors suggested this may indicate a lack of support after initial victimisation, or possible mutual experiences of sexual extortion in relationships between minors.

## Challenges and opportunities for research into sexual extortion of minors

While each of the studies detailed significant impacts experienced by victims, it is important to acknowledge the challenges in studying sexual extortion against minors. Children can be a difficult population to access and study for many reasons, including ethical concerns, especially when dealing with sensitive topics such as sexual extortion. To overcome this challenge, some of the included studies relied on open-source information (eg news articles), interviews with participants who served as proxies for victims (eg school counsellors), or official or informal reports of victimisation. While contributing to the knowledge area, these data sources may introduce biases emerging from small sample sizes or from a reliance on third parties rather than the victims themselves. Alternatively, many studies depended on self-report data through surveys. While surveys may be limited by poor recall and impression management (Chan 2010), they can nevertheless allow researchers to draw representative samples and collect specific information about the topic of interest.

Challenges also arise when defining what constitutes sexual extortion of a minor. When researching an emerging topic such as this, conceptual and operational definitions can vary between studies. While we limited the systematic review to studies fitting certain definitional parameters, some differences were still apparent, such as studies focusing only on online sexual extortion or financial extortion. While allowing for insight into specific subject matter, these approaches can introduce challenges with cross-study comparisons and the development of a cohesive field of research.

Although secondary sources such as news articles or interviews with proxies for victims can provide insight into sexual extortion, is important that, where possible, opportunities are taken to conduct research using primary sources. These include interview or survey data relating to victims or perpetrators of sexual extortion. Further, it is pivotal that consistent definitions are adopted to allow development of a comparable field of knowledge. Together, these opportunities offer a pathway to strengthening the growing knowledge base on sexual extortion of minors, and in turn informing prevention, intervention and support approaches.

## Discussion

Minors who were victims of sexual extortion experienced a range of psychological, social-environmental, financial and offending or criminal justice impacts both immediately following the extortion and for prolonged periods after. Victims also experienced impacts related to ongoing and poly-victimisation, including physical assaults, stalking, or threats of assault. The psychological impacts of sexual extortion victimisation were the most common reported in the literature reviewed, and included poor mental wellbeing, distress and anxiety, fear, shame, humiliation, hopelessness, anger, self-blame, self-harm and, in some cases, attempted or actual suicide. Victims also commonly reported social isolation, and some had their schooling, employment and living situation disrupted. Financial costs were also incurred by victims, relating to both the extortion of money and the financial burden of help-seeking associated with victimisation. Importantly, some victims reported that police considered taking action against them for the production or dissemination of CSAM, while for others victimisation was associated with the development of offending behaviours.

These impacts should not be viewed in isolation; it is unlikely that a victim of sexual extortion would experience only one of these impacts. Rather, it is more likely that each harm may exacerbate or increase the likelihood of another. For example, social isolation may enhance the impact of victimisation on mental wellbeing, and the financial burden of extortion may lead to or enhance the psychological or social impacts on victims. The research included in this review typically sought to understand what types of impacts were associated with experiences of sexual extortion among minors. Future research should consider how impacts interrelate and connect, to build a more accurate picture of the harms produced by sexual extortion.

Future research should also consider the context in which sexual extortion is experienced, and whether this influences the severity of the impacts. Broader research has found that many instances of sexual extortion of a minor were perpetrated by an intimate partner (57%, Finkelhor, Turner & Colburn 2023; 31.9%, Patchin & Hinduja 2020; 64.2%, Wolak & Finkelhor 2016), and such cases of sexual extortion may form part of a pattern of abuse including other forms of simultaneous harm. These scenarios differ considerably from those involving online predators or profit-motivated cybercrime. Importantly, when sexual exploitation is experienced by minors online, it may influence the likelihood of revictimisation later in life (Teunissen et al. 2024). Understanding these differences and how they relate to the impacts of sexual extortion on victims is an important area for future research.

### Implications

The varied short- and long-term impacts of sexual extortion experienced by minors emphasises the importance of diverse and holistic prevention activities. Importantly, there are key differences between minors and adults in how sexual extortion is experienced, as well as how and why it occurs (Wolak et al. 2018), highlighting the need for prevention and support approaches to be tailored to and targeted towards children. Further, children as young as eight years old reported experiencing sexual extortion (NCMEC 2016). It is therefore pivotal that responses are age appropriate and relevant to the experiences of minors. Responses may include the provision of age-appropriate online safety and sexual education for children or their parents. An education approach could help equip children and young people with knowledge on risky online behaviours, and where to seek help, publicising programs that offer early intervention.

Improvements to help-seeking and reporting mechanisms are also important. Prior research indicates that many minors do not report sexual extortion (Wolak & Finkelhor 2016), and these findings emphasise that sexual extortion can be a prolonged experience without intervention (Açar 2016; Alsoubai et al. 2022; Wolak et al. 2018). Help-seeking is crucial to prevent victims from complying with demands, which in turn may help reduce their vulnerability to repeat victimisation (Thorn 2019; Wolak et al. 2018), while also providing support that can mitigate the impacts on wellbeing. To improve this, greater public awareness and accessibility of reporting avenues is needed (eSafety Commissioner 2019). Further, anonymous reporting may help improve bystander reporting of harmful online behaviour by reducing fear of repercussions (eSafety Commissioner 2019).

Proactive action could also be taken to prevent sexual extortion through initiatives such as Safety by Design, which redistributes the responsibility for protecting children online to service providers (eSafety Commissioner 2019). This is particularly relevant to sexual extortion as it often occurs on popular social media platforms (Canadian Centre for Child Protection 2022). Prevention approaches using Safety by Design principles may include ensuring accounts are set to private by default; introducing age and identity verification; and using artificial intelligence to detect, flag and remove harmful content (Meurens et al. 2022). Blocking malicious actors from being able to easily contact children online is an important measure in reducing perpetration and subsequent harms.

Sexual extortion of minors is an emerging issue with a growing body of research. Given the speed at which sexual extortion is proliferating and online spaces are developing, future research should continue to measure the prevalence of sexual extortion of minors, alongside changes to the impacts they experience, to help inform prevention and intervention initiatives and enhance the support available to victims.

## Acknowledgements

The authors give special thanks to Megan Whittle from the AIC's JV Barry Library for supporting us to conduct literature searches.

## References

*URLs correct as at September 2024*

*Sources included in review

Açar KV 2016. Sexual Extortion of Children in Cyberspace. *International Journal of Cyber Criminology* 10(2): 110–126. https://doi.org/10.5281/zenodo.163398

*Alsoubai A et al. 2022. From 'friends with benefits' to 'sextortion:' A nuanced investigation of adolescents' online sexual risk experiences. *Proceedings of the ACM on Human-Computer Interaction* 6(CSCW2): 1–32. https://doi.org/10.1145/3555136

Anderson M, Faverio M & Gottfried J 2023. *Teens, social media and technology 2023*. Washington, DC: Pew Research Center. https://www.pewresearch.org/internet/2023/12/11/teens-social-media-and-technology-2023/

*AUSTRAC 2022. *Combating the sexual exploitation of children for financial gain: Financial crime guide*. AUSTRAC. https://www.austrac.gov.au/business/how-comply-guidance-and-resources/guidance-resources/combating-sexual-exploitation-children-financial-gain

Australian Centre to Counter Child Exploitation 2023. *Sextortion: Messaging kit*. https://www.accce.gov.au/sites/default/files/2023-11/Sextortion_messaging_kit_October2023.pdf

Canadian Centre for Child Protection 2022. An analysis of financial sextortion victim posts published on r/sextortion. https://protectchildren.ca/en/resources-research/an-analysis-of-financial-sextortion-victim-posts-published-on-sextortion/

Canadian Centre for Child Protection 2017. *Survivors' Survey: Full report 2017*. Winnipeg: Canadian Centre for Child Protection. https://protectchildren.ca/en/resources-research/survivors-survey-results/

Chan D 2010. So why ask me? Are self-report data really that bad? In CE Lance & RJ Vanderberg (eds), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in organizational and social sciences* 1st ed. Routledge: 329–356. https://doi.org/10.4324/9780203867266-22

Chiang E & Grant T 2019. Deceptive identity performance: Offender moves and multiple identities in online child abuse conversations. *Applied Linguistics* 40(4): 675–698. https://doi.org/10.1093/applin/amy007

Cross C, Holt K & O'Malley RL 2022. "If u don't pay they will share the pics": Exploring sextortion in the context of romance fraud. *Victims & Offenders* 18(7): 1–22. https://doi.org/10.1080/15564886.2022.2075064

*Dolev-Cohen M 2023. A qualitative study of how Israeli Arab educational counselors handle image-based sexual abuse of female adolescents. *Violence Against Women*. Advance online publication. https://doi.org/10.1177/10778012231179219

*Dolev-Cohen M, Nezer I & Zumt AA 2023. A qualitative examination of school counselors' experiences of sextortion cases of female students in Israel. *Sexual Abuse* 35(8): 903–926. https://doi.org/10.1177/10790632221145925

Douglass CH, Wright CJC, Davis AC & Lim MSC 2018. Correlates of in-person and technology-facilitated sexual harassment from an online survey among young Australians. *Sexual Health* 15(4): 361–365. https://doi.org/10.1071/SH17208

Eaton AA, Ramjee D & Saunders JF 2023. The relationship between sextortion during COVID-19 and pre-pandemic intimate partner violence: A large study of victimization among diverse U.S men and women. *Victims & Offenders* 18(2): 338–355. https://doi.org/10.1080/15564886.2021.2022057

eSafety Commissioner 2019. Safety by design overview. Australian Government. https://www.esafety.gov.au/industry/safety-by-design

Europol 2015. *The Internet Organised Crime Threat Assessment (IOCTA) 2015*. https://www.europol.europa.eu/publications-events/main-reports/internet-organised-crime-threat-assessment-iocta-2015

*Finkelhor D, Turner H & Colburn D 2023. Which dynamics make online child sexual abuse and cyberstalking more emotionally impactful: Perpetrator identity and images? *Child Abuse & Neglect* 137: 106020. https://doi.org/10.1016/j.chiabu.2023.106020

Henry N, Flynn A & Powell A 2019. Responding to 'revenge pornography': Prevalence, nature and impacts. Report to the Criminology Research Advisory Council. https://www.aic.gov.au/crg/reports/crg-0815-16

Kopecký K 2017. Online blackmail of Czech children focused on so-called "sextortion" (analysis of culprit and victim behaviours). *Telematics and Informatics* 34(1): 11–19. https://doi.org/10.1016/j.tele.2016.04.004

Liggett R 2019. Exploring online sextortion. *Family & Intimate Partner Violence Quarterly* 11(4): 45–56

Meurens N, Notté E, Wanat A & Mariano L 2022. *Child safety by design that works against online sexual exploitation of children.* Down to Zero Alliance. https://doi.org/10.13140/RG.2.2.15152.23043

Moher D et al. 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine* 151(4): 264–269. https://doi.org/10.1371/journal.pmed.1000097

National Center for Missing & Exploited Children 2023. *CyberTipline 2022 report*. https://www.missingkids.org/content/dam/missingkids/pdfs/2022-CyberTipline-Report.pdf

*National Center for Missing & Exploited Children 2016. *Trends identified in CyberTipline sextortion reports*. https://www.missingkids.org/content/dam/missingkids/pdfs/ncmec-analysis/sextortionfactsheet.pdf

NCMEC—see National Center for Missing & Exploited Children

*Nilsson MG, Tzani-Pepelasis C, Ioannou DM & Lester D 2019. Understanding the link between sextortion and suicide. *International Journal of Cyber Criminology* 13(1): 55–69. https://doi.org/10.5281/zenodo.3402357

*O'Malley RL & Holt KM 2022. Cyber sextortion: An exploratory analysis of different perpetrators engaging in a similar crime. *Journal of Interpersonal Violence* 37(1–2): 258–283. https://doi.org/10.1177/0886260520909186

*Patchin JW & Hinduja S 2020. Sextortion among adolescents: Results From a national survey of U.S. youth. *Sexual Abuse* 32(1): 30–54. https://doi.org/10.1177/1079063218800469

Salter M et al. 2021. Production and distribution of child sexual abuse material by parental figures. *Trends & issues in crime and criminal justice* no. 616. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti04916

Teunissen C, Cahill M, Napier S, Cubitt T, Boxall H & Brown R 2024. Sexual exploitation of children on dating platforms and experiences of revictimisation as an adult. *Trends & issues in crime and criminal justice* no. 697. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77482

*Thorn 2019. *Sextortion: Summary findings from a 2017 survey of 2,097 survivors*. Thorn. https://www.thorn.org/wp-content/uploads/2019/12/Sextortion_Wave2Report_121919.pdf

Wittes B, Poplin C, Jurecic Q & Spera C 2016. Sextortion: Cybersecurity, teenagers, and remote sexual assault. *Brookings.* https://www.brookings.edu/articles/sextortion-cybersecurity-teenagers-and-remote-sexual-assault/

*Wolak J & Finkelhor D 2016. *Sextortion: Findings from an online survey about threats to expose sexual images*. Crimes Against Children Research Center and Thorn. https://www.unh.edu/ccrc/sites/default/files/media/2023-05/sextortion-report-final-june-2016.pdf

*Wolak J, Finkelhor D, Walsh W & Treitman L 2018. Sextortion of minors: Characteristics and dynamics. *Journal of Adolescent Health* 62(1): 72–79. https://doi.org/10.1016/j.jadohealth.2017.08.014

This work was originally published as *Trends & issues in crime and criminal justice* no. 710. It is available online at **https://doi.org/10.52922/ti77789**.

# 12. Underage use of adult-based platforms: Risks and experiences

Tahlia Hart

## Introduction

A recent marketplace study suggests around 413 million individuals worldwide use dating apps (Market.us 2024). Dating apps and websites typically require users to be 18 years or over (Samuelson 2018). However, research suggests not all users are adults (Nagata et al. 2024; Thorn 2023). Recently, the Australian Institute of Criminology found 8.8 percent (*n*=883) of surveyed participants (*n*=9,987) had used a dating app while underage (Teunissen et al. 2024). Children as young as 11 have been found to use dating apps (Nagata et al. 2024). Adult users of these platforms engage in various activities with their matches, both online and offline; however, the experiences of children using these services are less well known. This article presents findings on the online and offline behaviours of young people who, as children, used an adult-based platform (ABP)—that is, dating apps or websites, remunerative matchmaking platforms (eg Seeking, Sugar Daddy), subscription-based content-sharing platforms (eg OnlyFans) and online fetish communities (eg FetLife).

### Underage use of ABPs

Understanding children's use of ABPs will provide insight into how they use platforms intended for adults and the risk-taking opportunities, if any, that result from this. Cooper et al. (2023) suggest that children use dating apps as a playful social activity with friends, which may mean children use ABPs for enjoyment. Evidence shows that it is developmentally normal for children to enjoy risk-taking behaviours (Posick 2020). Accordingly, children may seek out, and engage in, risk-taking activities on ABPs (eg Mignault et al. 2022). It is recognised that individuals may take such risks on and off ABPs, given that these platforms provide online and offline connection opportunities (Phan, Seigfried-Spellar & Choo 2021; Sautter, Tippett & Morgan 2010). A summary of the existing literature about risk-taking on ABPs is below.

#### *Online experiences*

ABPs enable users to select and connect with matches through in-app features such as chats and video (Ward 2016). Studies suggest that adult users engage in risky online behaviours on these platforms, such as exchanging personal details (eg phone number, address) shortly after matching (Fansher & McCarns 2019). Children's communication with matches, either other children or adults, requires further exploration to determine whether it exposes children to risks.

The existing literature has minimally addressed the activities that children take part in on ABPs and with other users. Research suggests that children may communicate with multiple matches simultaneously (Mignault et al. 2022). Children's conversations on ABPs may include flirting (Lykens et al. 2019) or sexting (Baltz & Velten 2016). The Australian Institute of Criminology found that underage dating app users received requests from other users for sexual content such as videos or photos and, in some instances, were offered payment for the content (Teunissen et al. 2024). Another recent study suggests some underage dating app users were asked for nude content or to attend a hotel for sexual activity (Child Welfare League Foundation 2024). These studies provide minimal insight into underage experiences of ABPs.

### *Offline experiences*

The existing literature about offline risk-taking behaviours among children using ABPs is also limited. The activities children may seek out, or engage in, offline as a result of using ABPs are largely unknown. Existing studies suggest adult users of these platforms engage in offline risk-taking behaviours, including meeting offline shortly after matching, having unprotected sex, or using illicit substances with their matches (Fansher & McCarns 2019; Phan, Seigfried-Spellar & Choo 2021).

Children also meet matches offline (Cooper et al. 2023; Jozsa et al. 2021); male-identifying children with same-sex preferences are significantly more likely to do so (Thunyapipat, In-iw & Manaboriboon 2023). The location where children, as opposed to adults, meet their matches is minimally understood. Sinno, Macapagal and Mustanski (2023) found that children who receive sexual education from their families are significantly more likely to meet users in public locations than those who did not receive it. Once a child meets a match offline, there is little understanding of the activities they engage in. Jozsa et al. (2021) found underage ABP users spend less time communicating on dating apps compared to social media platforms before meeting a match offline. Mignault et al. (2022) found about 20 percent of participants (children and young adults) met a match offline on the same day or the next day, often resulting in sex. Vandenbosch et al. (2016) found children's use of dating websites significantly predicted their sexual activity. Studies of children who use dating apps suggest that most who met matches offline 'hooked up' and engaged in sexual activities (Jozsa et al. 2021; Macapagal et al. 2019). Similarly, children who use remunerative matchmaking platforms may meet matches offline to fulfil a sexual and monetary agreement (Li 2015). These findings suggest that children may be engaging in risky sexual behaviours, possibly before the age of consent.

Macapagal et al. (2019) note that where children's online ABP matches became sexual partners, the matches were older than the child. This may constitute a criminal offence for an adult if a child is under the age of consent. Another study of 11- to 25-year-old participants found some (30.6%) had used substances (eg alcohol, vasodilator, methamphetamine, cannabis) during sexual intercourse with matches (Thunyapipat, In-iw & Manaboriboon 2023); however, it is unclear who provided the drugs. These findings indicate that children may be exposed to a myriad of behaviours that could be deemed criminal.

The above provides some insight into children's activities on ABPs. Yet little is known about how these children interact, the information they share, whether conversations move to other platforms or from online to offline, and the activities (criminal or otherwise) they engage in with their matches. The current study addresses these matters.

## Method

Australian participants aged 18 to 21 years of age ($n$=46) who used at least one ABP as a child were interviewed about their underage use. Participants were recruited via Meta (Facebook and Instagram) advertisements between June and November 2021. This recruitment method allowed for advertising to be targeted at Australians aged 18 to 21 years old. Individual semi-structured audiovisual interviews (approximately 30 to 90 minutes) were conducted via Microsoft Teams. Ethics approval (2904) was obtained from Flinders University's Human Research Ethics Committee.

Participants were located in New South Wales (*n*=14), Queensland (*n*=12), Victoria (*n*=9), South Australia (*n*=4), Western Australia (*n*=2), the Australian Capital Territory (*n*=2), the Northern Territory (*n*=2) and Tasmania (*n*=1). Participants self-identified as male (*n*=22), female (*n*=16) or another category (eg nonconforming, non-binary, genderfluid, unlabelled; *n*=8). Participants self-identified as heterosexual (*n*=19) or various LGBTQ+ categories (*n*=27).

Participants were asked questions about their use of ABPs in online and offline settings, including:

- Did you have a conversation with another user on an ABP?
- What types of conversations did you have?
- Did conversations ever move to another platform (eg SMS, Facebook)?
- Did you, or another user, request to meet offline?
- What was the purpose of meeting offline?
- Did you engage in activities that are illegal for a minor (eg alcohol/tobacco use) or generally (eg illicit substance use)?

The interviews were recorded, transcribed and uploaded to NVivo for analysis. The above questions are part of a broader study that considered a child's entire process of ABP use. An adaptive theory and crime script analysis approach was taken to code and analyse the data. Interview transcripts were individually reviewed, and the responses were initially coded into the broad categories of preparation, pre-activity, activity, and post-activity. The categories with participants' coded responses were then reviewed and further subcategorised into distinct themes. For example, the data in the activity category, which are relevant for underage experiences, were separated into experiences in online and offline environments before being subcategorised into similar experiences within those environments (eg alcohol use, conversation types). The similarities, differences and recurrent themes between the participants were noted down, as were the data similarities to existing theories and studies. The coded data were continually revisited to ensure appropriate coding. The responses are discussed below.

## Limitations

Participants were adults as opposed to children, so current insights about underage use of ABPs were not obtained. Adults were recruited due to the perceived difficulties associated with interviewing children about sensitive topics such as sex. The participants may have still provided socially desirable responses about sensitive topics or their transgressive behaviour (see Guest, Bunce & Johnson 2013). Evidence shows that participants may engage in sensitive research online if they remain anonymous (Mendelson 2007). In this study, participants were interviewed online and could disable their camera, which may ensure trust and comfort while alleviating anxiety (Deakin & Wakefield 2013). Participants in the present study appeared to freely discuss sensitive topics without hesitation.

The authenticity of participant perceptions may be affected if their perceptions at the time of interview differed from when they were underage on ABPs, possibly due to changes in their maturity (see Wood & Wheatcroft 2020). Further, adult recollections about their underage behaviour may be unreliable given the time that had passed; however, participants were asked to clarify inconsistent responses to minimise possible recall or self-report biases.

# Results

This study found that the participants, as children, engaged in a range of risk-taking activities in both online and offline environments arising from their use of ABPs. A summary of these findings is provided below.

## Online experiences

### *Communication with matches*

Almost all participants (*n*=45) as children conversed with one or more match(es) via messaging features built into ABPs or on other platforms. Many participants (*n*=38) shared their social media details with a match and moved their conversation to another online platform (eg Instagram, Snapchat, Facebook, Skype, Kik) or mode of communication such as SMS. Some participants (*n*=12) moved conversations restricted on ABPs (eg photo or video sharing) to another platform (eg Snapchat, Kik). Some participants, such as Participant 44 (trans woman, bisexual), identified that some platforms, like Snapchat, afforded additional privacy compared to ABPs: 'Snapchat is just the go-to app on Grindr for nudes because [the photo] obviously doesn't save.'

### *Types of communication*

Participants reported that as children they engaged in various types of conversations depending on their 'goal' with a match (eg casual or friendly (*n*=30), 'deep' (*n*=4), flirty (*n*=20), sexual (*n*=15)). It is expected that these conversation goals were fluid, given that some participants (*n*=7) identified flirting could be used to progress to sexual conversations.

Participants identified that sexual conversations included discussions about their past experiences or possible future encounters. About half of the participants (*n*=24) as children either shared or received sexually explicit photos or videos with a match via messages; however, most of the participants (*n*=29) had experienced sexual conversations or sending/receiving sexual content before using ABPs underage. Almost half of the participants (*n*=20) were comfortable sexting with their matches, and some (*n*=6) stated prior experiences made them confident to do so. Some participants refused to engage in sexting (*n*=17), whereas some did so reluctantly (*n*=6). Most participants (*n*=31) suggested that their prior experience of online conversations outside of ABPs assisted them with these conversations. For example, Participant 35 (female, bisexual) stated that a conversation on an ABP:

> … was no different than if someone messaged me on Facebook … The set-up for messages is basically the exact same, the only difference is … 'Oh, look, he's matched, you both think each other are cute'.

Children may not, therefore, be learning new communication skills on ABPs; however, some participants (*n*=4) used these platforms to learn how to flirt in both online and offline environments. Some participants (*n*=10) reported lacking the knowledge to take part in sexual conversations—for example, Participant 34 (non-binary, gay):

> I would try and act like I knew what they were talking about, and what I was talking about, so it was very much going along with what they were saying.

A small number of participants (*n*=6) also identified receiving validation from their matches, as well as from swiping and matching. Participant 15 (female, bisexual) explained:

> It's like, 'Mm yes, people want to flirt with me, people were giving me attention, look at me!' You just felt like you were on top of the world … it was just an endorphin rush.

Participant 34 (female, bisexual) expanded on this:

> The sheer volume of men trying to get your attention … you'll never ever get someone just coming up to you being like, 'I need your phone number, you are the most beautiful person I've ever seen', but on Tinder that was a daily thing.

Participants experienced validation when their matches initiated or had flirty/sexual conversations or sent compliments. Some participants (*n*=5) retrospectively recognised the validation they received was not what they wanted or that it was short-lived with ultimately negative connotations, as explained by Participant 15 (female, bisexual):

> I want to feel good about myself, you've had a shit day, and just getting complimented, especially by guys that you found attractive … It just was a rush … but then after five seconds you were like, 'Ew no'.

Participant 15 identified how thrills elicited through online environments could quickly turn into negative experiences while underage. Participant 31 (non-binary, queer) similarly explored this:

> I think part of what I wanted from [ABPs] was validation and a bit of self-worth at times. It's kind of hard to say if I actually got that or if it was just a quick gratification, but ultimately doing more harm.

## Offline experiences

### *Meeting offline*

Half of the participants (*n*=23) met an ABP match offline as children. Among those for whom information was available, some met with one match (*n*=7) and others with multiple matches (eg between two and 40; *n*=13). Participants reported various reasons for meeting offline (eg 'catching-up' *n*=9; sex *n*=13; dating *n*=10). Some participants met someone offline once (*n*=6) or multiple times (*n*=13); the latter was associated with relationships (*n*=10) or ongoing casual sex (*n*=3). Some participants (*n*=10) reported that, as children, they had previously met people offline from other platforms (eg Yubo, Whisper, Omegle, Kik) for similar reasons (eg friendship, sex, dating).

A participant's motivation appeared to influence where they met an ABP match offline. Participants tended to report meeting a match in a public place (eg restaurant, campus, mall, party, park, beach; *n*=17) rather than in a private place such as a house (*n*=6). Participants suggested they met publicly for safety (*n*=4). For example, Participant 35 (female, bisexual) stated, 'I never meet anyone off the internet in private, I think that's one sure-fire way to be murdered'. Participants who met in private reported sexual activity (*n*=6). Some participants (*n*=6) transitioned from public meet-ups to private, as explained by Participant 8 (female, queer): 'Publicly first, just to be like, "This is who we are" and then we went back to their place.'

While some participants reported meeting with other underage users (*n*=4), others reported meeting up with adult matches, aged between 18 and 40 years of age, when they were children (*n*=17). Participant 36 (male, gay) explained this:

> He was in his early forties and we ended up meeting up off the app and I was 14 years old, and he knew that.

Some participants (*n*=8) told their match their true age before meeting offline. Only a couple (*n*=2) reported that their disclosure prevented the meet-up from happening. Participant 8 (female, queer) explains this was uncommon: 'There might have been a couple that were cancelled because they found out that I wasn't 18, but it wouldn't have been too many.'

A few participants (*n*=3) were queried by matches about their age when meeting offline, including Participant 25 (male, heterosexual):

> … I still had quite bad acne across my forehead. They kind of said, 'Oh, you look like you're still in high school' …

*Sexual behaviours*

Some participants (*n*=13) identified that they had sex with either an underage or adult match as a child. Participant 23 (male, heterosexual) explained that this was his intention:

> People who you didn't know rocked up [to a party] and you met new people … There were various sexual encounters and whatnot: the majority of the time I guess the goal of [my ABP use] was a success.

The participants (*n*=7) rarely identified sexual activities with adults to be risky. Only one participant, who used a remunerative matchmaking platform underage (*n*=9), engaged in sexual activities offline for money. Participant 37 (female, heterosexual) was 14 years of age and now regrets it:

> There were three different times that I met up with someone … he wanted to meet up and offer me $700 each time and we actually did have sex … Pretty much being a prostitute, pretty nasty.

*Criminal activity*

Some participants (*n*=13) reported engaging in criminal activities with a match offline while underage (eg alcohol use *n*=10; drug use *n*=5; tobacco use *n*=1; breaking into private premises *n*=1). Participants suggested the match supplied the drugs (*n*=3) or alcohol (*n*=6). Some participants, like Participant 2 (nonconforming, queer), were intentional in this:

> It was my friend's birthday, and we were both underage and I wanted to get alcohol for her … So I went out with that person with the promise of sexual activity in exchange for them purchasing alcohol for me.

Many of the participants were already engaging in these activities outside of ABPs while underage. Some participants (*n*=9) identified that they used a fake ID to access age-restricted services; however, a larger number (*n*=32) suggested they were able to access these services without a fake ID by lying about their age, appearing older or asking peers, or because ID requirements were ignored (eg licensed venues, tobacco, alcohol, tattoos, piercings, R18+ games).

# Discussion

## Online experiences

*Communication with matches*

As noted above, participants engaged in various types of conversations on and off ABPs with their matches while underage. Participants were found to communicate with numerous matches concurrently (Newett, Churchill & Robards 2018). Accordingly, these connections may provide opportunities for children to take risks.

Participants were also found to move their conversations from ABPs which restricted communication (eg not allowing photos or videos to be shared) to other online platforms that allowed those features (Licoppe 2020). Participants may have moved from ABPs to Snapchat to engage in sexting because it is a popular platform for this activity (Flannery 2016). As noted above by Participant 44, there is a misconception that Snapchat content is not risky because images are deleted; however, users can save, screenshot or record content with a device (Mehan 2022).

Sexting between children and adults can be illegal given the age of sexual consent (Baltz & Velten 2016; Macapagal et al. 2018). Children may experience feelings of validation and status by sexting (Giambonini 2021). Adults using ABPs may also knowingly groom children through sexting; this practice may elicit feelings of excitement and fear from a child (Whittle, Hamilton-Giachritsis & Beech 2014).

*Types of communication*

The types of communication used by participants on ABPs were diverse, which is unsurprising given that interactions are generally dynamic (Fitzpatrick & Birnholtz 2018). However, it is also noted that grooming techniques can resemble all conversation types (eg casual, secret, flattery, sexual, kind, nasty, erratic, familiar, trusting; Whittle, Hamilton-Giachritsis & Beech 2014). Further, underage ABP users may be groomed without their knowledge, given that children may not understand who they are communicating with and the associated risks (Karklins & Dalton 2012; Wood & Wheatcroft 2020). In this study, underage users appeared to communicate based on their intentions (eg 'deep' conversations were often associated with developing a further connection for dating; Ward 2016). Accordingly, not all of children's communications on ABPs are 'risky'; however, there are instances of risk-taking behaviour arising from conversations (see further below). Posick (2020) suggests a child's prefrontal cortex, which assists in making careful decisions, is not fully developed until they reach their twenties. Children are therefore more likely than adults to seek out instant rewards and underestimate the possible risks (Posick 2020). Children on ABPs, while focusing on rewards, may unknowingly engage with adult users who intend to groom them either online or offline.

This study found that children may take part in flirty and sexual conversations with matches; however, they were sometimes having similar conversations on child-based platforms before using ABPs. It is suggested in the literature that children enjoy sexting if they remain comfortable with their actions (Hart 2017). Lykens et al. (2019) suggest children will confidently initiate conversations in online environments by flirting. Young people may regret sexting with older people because of the retrospective feelings of guilt or shame (Bloom et al. 2022). Accordingly, while the findings suggest that ABPs are unlikely to be the initial instigators of childhood risk-taking behaviour, these platforms allow children to engage in risky communications.

The findings also suggest that children already had the communication skills needed to engage in conversations on ABPs; however, children can also use ABPs to learn or to enhance their communication skills in areas they are unsure about or inexperienced in (eg flirting, sexual conversations). Children may be exposed to risk when engaging in these types of conversations, especially if they communicate with adults intentionally seeking children for sexual activities.

As noted above, children's experiences of excitement are not always rewarding, particularly when the excitement develops into anxiety or fear (Pimentel & Steinmetz 2022). Some participants noted feeling validated through their underage use of ABPs; however, their enjoyment was short-lived if they later experienced guilt or felt sexually objectified in the process (Baltz & Velten 2016). Another study of adolescent male-seeking-male ABP users found that underage users felt sexualised and objectified when they encountered another user who was older or 'creepy', who requested sexual content or who outright rejected the concept of a relationship (Owens et al. 2021).

## Offline experiences

### *Meeting offline and sexual behaviours*

Underage ABP users met matches offline; however, the participants in this study appeared to have met with fewer matches on average than those in other studies (Macapagal et al. 2019). This finding confirms that children do not restrict their ABP matches to online conversations but also extend their connections to offline environments, which increases their chances of engaging in risk-taking behaviours. The purpose of meeting a match offline differed (eg 'catching up', 'hooking up'). This is unsurprising given adolescence involves the exploration and commencement of sexual relationships (Valkenburg & Piotrowski 2017). Therefore, children may also be using ABPs to form legitimate connections.

Adolescents have been previously found to use 'common sense' approaches to protecting themselves (Torrijos-Fincias et al. 2021). Children may experience risks at any stage of their ABP use; however, research suggests that risky experiences typically occur within the initial communication stages and primarily to female-identifying users (Adeyinka-Skold 2020). The present study found underage ABP users learned or applied previously learned skills to protect themselves in online and offline environments. Participants' chosen locations for offline meetings appeared to be influenced by their purpose. Offline environments offer none of the protections or safety features of online environments such as blocking (Fitzpatrick & Birnholtz 2018). Underage ABP users in this study appeared to implement safety techniques such as meeting matches publicly, which they believed mitigated potential risks. Individuals may experience a fear of the unknown, risk of being deceived, or risk to their safety when meeting someone offline that they originally met online (Chambers 2021). Despite a child's attempt to be safe, it may be an inadequate response to the situation and may put them at risk of harm. Children's feelings of invincibility may also increase their tendency to engage in risk-taking behaviours (Valkenburg & Piotrowski 2017).

The age of an underage ABP user's match was unable to be independently verified from the interviews. Some participants reported they met with adults aged between 18 and 40 years, sometimes when they were under the age of consent. Previous studies found an underage user's matches were typically older than them by two or more years, which is unsurprising because ABPs are intended for adult audiences (Macapagal et al. 2019; Macapagal et al. 2021). Sexual activities between adults and children under the age of consent are illegal regardless of whether the parties express consent or conceal their age (Kuate-Defo 2004). Lippert et al. (2020) found adolescents experience an increased risk of coercion when communicating with people aged 21 years or older, possibly because adolescent minds are easy to manipulate (Olugbenga 2011). Previous research suggests underage ABP users who willingly engage sexually with adults may not consider themselves victims, but their perceived 'consent' does not discount the developmental differences that exist between adults and children (Macapagal et al. 2019). This study suggests underage users of ABPs are at risk of sexual predation, regardless of whether they can and do 'consent' to sexual activities. In addition to this, some participants used remunerative matchmaking platforms, which may create opportunities for child exploitation (Kuate-Defo 2004) and child sex trafficking (McKim 2021).

Participants reported that their adult matches did not readily query their true age. In the few instances when an adult was aware, they continued to meet offline, which suggests age may not be important to these matches or that they intentionally sought underage users. In the few instances that a child's age was queried, it was unlikely to prevent the offline meeting from happening. These findings are concerning; however, further insight into adult users of these platforms is required.

### *Criminal activity*

Adult users of ABPs who did not query a child's age when meeting offline could be linked to criminal behaviour. As explained above, sexual activities between an adult and child (under the age of consent) are illegal (Kuate-Defo 2004). This study's findings suggest that some participants may have been victims of statutory rape, depending on their age and jurisdiction.

This study found that some underage ABP users engage in criminal behaviour as a result of using these platforms, which is supported by Bocheński (2022), who found children and adults meet offline to consume alcohol. Adults who commit crimes against children may use alcohol as a way to encourage their victims to accept the abuse (Ringenberg et al. 2022). Another study about underage male-seeking-male ABP users found that most participants did not use alcohol (83%), marijuana (79%), or other drugs (92%) before having sex with their matches (Macapagal et al. 2021). The current study is unable to state whether drugs and/or alcohol were used by adult matches as either a prerequisite for inappropriate sexual activity with a child or as a leisure activity during their meet-up. Regardless, the present study suggests that ABPs allow underage users to engage in criminal activities; however, children can take part in these activities regardless. Similarly, Kypri et al. (2005) suggest adolescents access age-restricted products through illegitimate purchasing attempts or their peers.

## Future research

Future studies could include child participants to determine the risk-taking behaviours they currently engage in on ABPs. A study could assess the in-app features of ABPs that children are currently exposed to and the risk-taking opportunities they provide. Studies should consider surveying Australian children to discover the frequency of underage ABP use, as well as the age of onset. This may inform the prevalence of risk and at what age to target policy initiatives. Research addressing harm mitigation and preventative measures could assist in lessening children's access to these platforms.

# Conclusion

This study has demonstrated the various ways adults have reported engaging in risk-taking behaviours as children as a result of their using ABPs. This study has found that underage use of ABPs exposes children to risk not only in online environments but also in offline environments. The findings suggest that many of the activities participants engaged in while underage were activities they had previously experienced before or during their ABP use. Accordingly, ABPs are unlikely to be a cause of risk-taking behaviours but rather give adolescents the opportunity to continue to engage in these behaviours.

## Acknowledgements

The author acknowledges the participants involved in this research and appreciates their knowledgeable insight and time. This research was part of a PhD thesis for the Australian Research Council Discovery Project 'Becoming Delinquent Online: A Longitudinal Adolescent Study' (Project ID: DP170103538). The author thanks the PhD supervisors Emeritus Professor Andrew Goldsmith, Associate Professor Russell Brewer and Dr Melissa de Vel-Palumbo, as well as the independent reviewer Dr Russell Smith and Associate Professor Caitlin Hughes, for their ongoing support and feedback.

## References

*URLs correct as at May 2025*

Adeyinka-Skold S 2020. Barriers in women's romantic partner search in the digital age. In R Kalish (ed), *Young adult sexuality in the digital age*. Hershey PA: IGI Global: 113–137. https://doi.org/10.4018/978-1-7998-3187-7.ch007

Baltz H & Velten JC 2016. The influence of online dating on female adolescent maturation. *Journal of Education & Social Policy* 3(2): 15–22

Bloom BE, Kieu TK, Wagman JA, Ulloa EC & Reed E 2022. Responsiveness of sex education to the needs of LGBTQ+ undergraduate students and its influence on sexual violence and harassment experiences. *American Journal of Sexuality Education* 17(3): 368–399. https://doi.org/10.1080/15546128.2022.2033662

Bocheński M 2022. Is it pedophilia or not? Observations on child sexual abuse in Poland from a criminological perspective. *Crime Prevention and Community Safety* 24: 132–151. https://doi.org/10.1057/s41300-022-00141-5

Chambers D 2021. Dating app logic and geo-enabled mobile socialities. In A Hill, M Andersson & M Hartmann (eds), *The Routledge handbook of mobile socialities*. New York NY: Taylor & Francis: 101–114. https://doi.org/10.4324/9781003089872

Child Welfare League Foundation 2024. *2018 Survey on dating apps for children and teenagers and current situation of online dating leaving home*. https://www.children.org.tw/publication_research/research_report/2232

Cooper SC et al. 2023. A qualitative exploration of young australians' lived experiences of social media use and sexual agency. *Sexuality & Culture* 28: 534–553. https://doi.org/10.1007/s12119-023-10131-w

Deakin H & Wakefield K 2013. Skype interviewing: Reflections of two PhD researchers. *Qualitative Research* 14(5): 603–616. https://doi.org/10.1177/1468794113488126

Fansher AK & McCarns K 2019. Risky online dating behaviors and their potential for victimization. *Crime Victims' Institute* 4. https://doi.org/10.13140/RG.2.2.24083.84007

Fitzpatrick C & Birnholtz J 2018. "I shut the door": Interactions, tensions, and negotiations from a location-based social app. *New Media & Society* 20(7): 2469–2488. https://doi.org/10.1177/1461444817725064

Flannery M 2016. 'Tinder for teenagers' dating app being used by schoolchildren 'to send nude photographs'. *Mirror*, 30 September. https://www.mirror.co.uk/news/uk-news/tinder-teenagers-dating-app-being-8946294

Giambonini C 2021. *The meaning of sexting for teenagers: A social representational, post-Jungian investigation* (Doctoral thesis). Colchester: University of Essex

Guest G, Bunce A & Johnson L 2016. How many interviews are enough? *Field Methods* 18(1): 59–82 https://doi.org/10.1177/1525822x05279903

Hart M 2017. Being naked on the internet: Young people's selfies as intimate edgework. *Journal of Youth Studies* 20(3): 301–315. https://doi.org/10.1080/13676261.2016.1212164

Jozsa K et al. 2021. "Safe behind my screen": Adolescent sexual minority males' perceptions of safety and trustworthiness on geosocial and social networking apps. *Archives of Sexual Behaviour* 50: 2965–2980. https://doi.org/10.1007/s10508-021-01962-5

Karklins L & Dalton D 2012. Social networking sites and the dangers they pose to youth: Some Australian findings. *Current Issues in Criminal Justice* 24(2): 205–222. https://doi.org/10.1080/10345329.2012.12035955

Kuate-Defo B 2004. Young people's relationships with sugar daddies and sugar mummies: What do we know and what do we need to know? *African Journal of Reproductive Health* 8(2): 13–37. https://doi.org/10.2307/3583175

Kypri K, Dean J, Kirby S, Harris J & Kake T 2005. 'Think before you buy under-18s drink': Evaluation of a community alcohol intervention. *Drug and Alcohol Review* 24(1): 13–20. https://doi.org/10.1080/09595230500102731

Li JCM 2015. Adolescent compensated dating in Hong Kong: Choice, script, and dynamics. *International Journal of Offender Therapy and Comparative Criminology* 59(6): 588–610. https://doi.org/10.1177/0306624X13516285

Licoppe C 2020. Liquidity and attachment in the mobile hookup culture. A comparative study of contrasted interactional patterns in the main uses of Grindr and Tinder. *Journal of Cultural Economy* 13(1): 73–90. https://doi.org/10.1080/17530350.2019.1607530

Lippert T, Clary M, Bleoaja C, Walsh WA & Jones LM 2020. Statutory rape: Case characteristics when offenders are younger than 21 years of age versus older. *Journal of Interpersonal Violence* 37(9–10): NP7984–NP8005. https://doi.org/10.1177/0886260520975807

Lykens J et al. 2019. Google for sexual relationships: Mixed-methods study on digital flirting and online dating among adolescent youth and young adults. *JMIR Public Health and Surveillance* 5(2): e10695. https://doi.org/10.2196/10695

Macapagal K et al. 2021. Patterns of online and offline partnering, partnership characteristics, and condomless sex among adolescent sexual minority males in the USA. *AIDS and Behavior* 25(7): 2033–2045. https://doi.org/10.1007/s10461-020-03133-4

Macapagal K et al. 2018. Hookup app use, sexual behavior, and sexual health among adolescent men who have sex with men in the United States. *Journal of Adolescent Health* 62(6): 708–715. https://doi.org/10.1016/j.jadohealth.2018.01.001

Macapagal K, Kraus A, Moskowitz DA & Birnholtz J 2019. Geosocial networking application use, characteristics of app-met sexual partners, and sexual behavior among sexual and gender minority adolescents assigned male at birth. *Journal of Sex Research* 57(8): 1078–1087. https://doi.org/10.1080/00224499.2019.1698004

Market.us 2024. *Global online dating market, by platform (application, and web portals), by services (social dating, niche dating, matchmaking, and adult dating), by user (non-paying users, and paying users), by revenue generation, by region and companies: Industry segment outlook, market assessment, competition scenario, trends, and forecast 2023–2032*. United States: Market.us. https://market.us/report/online-dating-market/

McKim J 2021. How Grindr, a popular gay dating app, poses exploitation risk to minors. https://www.npr.org/2021/08/03/1024108203/how-grindr-a-popular-gay-dating-app-poses-exploitation-risk-to-minors

Mehan J 2022. *Artificial intelligence: Ethical, social and security impacts for the present and the future*. Cambridgeshire: IT Governance Ltd. https://doi.org/10.2307/j.ctv2k93td7

Mendelson C 2007. Recruiting participants for research from online communities. *Computers, Informatics, Nursing* 25(6): 317–323. https://doi.org/10.1097/01.NCN.0000299653.13777.51

Mignault L, Vaillancourt-Morel M, Ramos B, Brassard A & Daspe M 2022. Is swiping right risky? Dating app use, sexual satisfaction, and risky sexual behavior among adolescents and young adults. *Sexual and Relationship Therapy* 39(3): 819–842. https://doi.org/10.1080/14681994.2022.2078804

Nagata JM et al. 2024. Social epidemiology of online dating in U.S. early adolescents. *BMC Research Notes* 17(1): 145. https://doi.org/10.1186/s13104-024-06777-w

Newett L, Churchill B & Robards B 2018. Forming connections in the digital era: Tinder, a new tool in young Australian intimate life. *Journal of Sociology* 54(3): 346–361. https://doi.org/10.1177/1440783317728584

Olugbenga DO 2011. Adolescents and online dating attitudes. In E Dunkels, G Franberg & C Hallgren (eds), *Youth culture and net culture: Online social practices*. Hershey PA: IGI Global: 207–221. https://doi.org/10.4018/978-1-60960-209-3.ch012

Owens C et al. 2021. Motivations to start and stop using sexual networking applications among adolescent sexual minority men. *Sexuality Research and Social Policy* 20: 329–339 https://doi.org/10.1007/s13178-021-00641-3

Phan A, Seigfried-Spellar K & Choo KKR 2021. Threaten me softly: A review of potential dating app risks. *Computers in Human Behavior Reports* 3: 100055. https://doi.org/10.1016/j.chbr.2021.100055

Pimentel A & Steinmetz KF 2022. Enacting social engineering: The emotional experience of information security deception. *Crime, Law and Social Change* 77: 341–361. https://doi.org/10.1007/s10611-021-09993-8

Posick C 2020. Biosocial theories. In TJ Holt & AM Bossler (eds), *The Palgrave handbook of international cybercrime and cyberdeviance*. Cham: Springer: 705–721. https://doi.org/10.1007/978-3-319-78440-3_29

Ringenberg TR, Seigfried-Spellar KC, Rayz JM & Rogers MK 2022. A scoping review of child grooming strategies: Pre- and post-internet. *Child Abuse & Neglect* 123: 105392. https://doi.org/10.1016/j.chiabu.2021.105392

Samuelson K 2018. Gay male teens use adult hookup apps to find friends, partners. https://news.northwestern.edu/stories/2018/may/gay-male-teens-use-adult-hookup-apps-to-find-friends-partners/

Sautter JM, Tippett RM & Morgan SP 2010. The social demography of internet dating in the United States. *Social Science Quarterly* 91(2): 554–575. https://doi.org/10.1111/j.1540-6237.2010.00707.x

Sinno J, Macapagal K & Mustanski B 2023. Social media and online dating safety practices by adolescent sexual and gender diverse men: Mixed-methods findings from the SMART study. *Journal of Adolescent Health* 74(1): 113–122. https://doi.org/10.1016/j.jadohealth.2023.07.030

Teunissen C, Cahill M, Napier S, Cubitt T, Boxall H & Brown R 2024. Sexual exploitation of children on dating platforms and experiences of revictimisation as an adult. *Trends & issues in crime and criminal justice* no. 697. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77482

Thorn 2023. *LGBTQ+ youth perspectives: How LGBTQ+ youth are navigating exploration and risks of sexual exploitation online*. https://www.thorn.org/research/library/lgbtq-teens-are-at-a-greater-risk-for-exploitation-online/

Thunyapipat C, In-iw S & Manaboriboon B 2023. Influence of online dating apps on sexual risk behaviors among homosexual and bisexual adolescents and youths in Thailand: An online cross-sectional survey. *Siriraj Medical Journal* 75(9): 612–621. https://doi.org/10.33192/smj.v75i9.264171

Torrijos-Fincias P, Serrate-González S, Martín-Lucas J & Muñoz-Rodríguez JM 2021. Perception of risk in the use of technologies and social media: Implications for identity building during adolescence. *Education Sciences* 11(9): 523–534. https://doi.org/10.3390/educsci11090523

Valkenburg PM & Piotrowski JT 2017. *Plugged in: How media attract and affect youth*. New Haven CT: Yale University Press. https://doi.org/10.12987/yale/9780300218879.001.0001

Vandenbosch L, Beyens I, Vangeel L & Eggermont S 2016. Online communication predicts Belgian adolescents' initiation of romantic and sexual activity. *European Journal of Pediatrics* 175(4): 509–516. https://doi.org/10.1007/s00431-015-2666-6

Ward J 2016. Swiping, matching, chatting: Self-presentation and self-disclosure on mobile dating apps. *Human IT: Journal for Information Technology Studies as a Human Science* 13(2): 81–95

Whittle HC, Hamilton-Giachritsis CE & Beech AR 2014. "Under his spell": Victims' perspectives of being groomed online. *Social Sciences* 3(3): 404–426. https://doi.org/10.3390/socsci3030404

Wood A & Wheatcroft JM 2020. Young adult perceptions of internet communications and the grooming concept. *SAGE Open* 10(1). https://doi.org/10.1177/215824402091457

This work was originally published as *Trends & issues in crime and criminal justice* no. 722. It is available online at **https://doi.org/10.52922/ti78021**.

# 13. Victim-survivors' reflections on best practice in restorative justice for domestic, family and sexual violence

Siobhan Lawler

## Introduction

Restorative justice (RJ) is now integrated into criminal justice practice across Australia and internationally. RJ is a process that brings together people impacted by an offence (victim-survivors, perpetrators, other relevant members of the community) to take part in facilitated dialogue and problem solving to respond to the drivers and the impacts of the crime (Marshall 1999). In contrast to the courts, where victim-survivors are only peripherally involved in the ascertainment of guilt and sentencing, victim-survivors have a central role in RJ as active participants and decision-makers (Robinson & Shapland 2008).

Research with victim-survivors of domestic and family violence (DFV) and sexual violence who are navigating the criminal justice system shows that they often feel disrespected, silenced, misrepresented and excluded from the process (Herman 2005; Senate Legal and Constitutional Affairs References Committee 2023). Contrastingly, victim-survivors report satisfaction with RJ across a broad range of offences, ranging from minor crimes such as theft to very serious crimes including sexual assault, DFV and the murder of a loved one (Nascimento, Andrade & De Castro Rodrigues 2023; Sardina & Ackerman 2022). Benefits include lower anger, less shame and self-blame, and improved wellbeing (Silva 2020); reduced post-traumatic stress symptoms such as avoidance and intrusive thoughts; and reduced anxiety, fear and distress (Nascimento, Andrade & De Castro Rodrigues 2023).

Despite considerable evidence for, and interest in, RJ as a facilitator of healing and recovery from crime, there is a level of apprehension within the community and among some victim-survivors of sexual violence about whether and how RJ should be an option for these crimes (Jülich 2006; Marsh & Wagner 2015). There are concerns about the unique risks associated with DFV and sexual violence matters, particularly regarding the potential for perpetrators to use coercion and to manipulate the process (Gavrielides 2015; Pennell et al. 2021). Any expectation of apologies or forgiveness is problematic in the context of DFV relationships, where perpetrators frequently use apologies as a means of control, and forgiveness may not be in victim-survivors' interests (Stubbs 2007). However, other stakeholders have suggested that the challenges can be mitigated when programs adhere to best practice principles in design and implementation (Jeffries, Wood & Russell 2021).

Research with practitioners and victim-survivors has informed the development of best practice principles for RJ in DFV and sexual violence matters (Bolitho 2015; Gavrielides 2015; Jeffries, Wood & Russell 2021; Wolthuis & Lünnemann 2016). These include:

- specialist expertise in the dynamics and impacts of DFV and sexual violence;
- flexibility in approach to preparation and facilitation;
- an intensive preparation phase allowing for rigorous, continuous risk assessment;
- assessment of the victim-survivor's justice needs and the offender's capacity (or willingness) to meet these, and effective management of participant expectations around this;
- close engagement and collaboration with relevant services and experts; and
- a co-Convenor model where two Convenors are assigned to each matter, ideally representative of gender.

Stakeholders agree that best practice in RJ for DFV and sexual violence is characterised by a victim-centred process. Victim-centred RJ is distinct from other forms of RJ that can proceed without victim-survivor involvement (eg some youth justice conferencing models), or RJ processes that are not driven primarily by the decision-making, needs and interests of the victim-survivor (Jeffries, Wood & Russell 2021; Joudo Larsen 2014). Daly (2006) summarised the potential benefits of RJ for victim-survivors of gendered violence, including participation and voice, validation, offender accountability, relationship repair, and access to flexible, trauma-informed justice that can be adapted to their needs. Interviews with participants of RJ programs show that addressing justice needs, Convenor skills, extensive risk assessment and involvement of a range of experts are important to victim-survivors and key for safeguarding their safety (Bolitho 2015; Gavrielides 2015).

Prior research on the topic of best practice in RJ for DFV and sexual violence has focused on understanding the benefits victim-survivors may experience through RJ and what stakeholders consider important for safe practice (Jeffries, Wood & Russell 2021). What is missing from the literature is an analysis of Australian victim-survivors' experiences of RJ juxtaposed against the best practice principles for RJ and DFV and sexual violence discussed in the literature. This is important for developing understanding of victim-survivors' experiences of best practice in Australia and internationally, and building upon evidence primarily informed by the opinions of stakeholders.

This study examines the experiences of victim-survivors participating in an RJ program for DFV and sexual violence, canvassing what they viewed positively about this program and any areas for improvement they identified. This is used to explore best practice in RJ for DFV and sexual violence from the perspective of victim-survivors. The analysis and subsequent discussion examine their views against best practice principles identified in the literature, both adding to and expanding on them.

## The Restorative Justice Scheme

This study is part of a larger research project evaluating Phase Three of the ACT Restorative Justice Scheme for DFV and sexual violence matters. In 2019 the Australian Institute of Criminology was commissioned by the ACT Government Justice and Community Safety Directorate to evaluate the processes and outcomes of this scheme. The broader project involved interviews with key stakeholders and participants, analysis of post-conference surveys, observation of conferences and analysis of administrative data to map referral pathways and recidivism outcomes in matters referred between November 2018 and August 2022 (see Lawler, Boxall Dowling 2025 for further details).

Restorative justice in the Australian Capital Territory is delivered through the Restorative Justice Scheme (RJS) and legislated under the *Crimes (Restorative Justice) Act 2004* (ACT). The RJS has been taking referrals from the criminal justice system for less serious offences by young people since 2005 (Phase One). The RJS was expanded in 2016 to include adults and more serious offences by young people and adults (Phase Two). In 2018 the RJS commenced taking referrals for DFV and sexual violence matters (Phase Three).

Consistent with best practice principles of RJ for DFV and sexual violence (eg Jeffries, Wood & Russell 2021; Wolthuis & Lünnemann 2016), Phase Three includes a number of unique features which distinguish it from Phases One and Two, including:

- a co-Convenor model (where two Convenors are assigned to each case);
- more intensive risk assessment and case review processes by a member of the senior leadership team; and
- careful consideration of any patterns of abusive behaviour and histories of harm in the relationship.

DFV and sexual violence matters can be referred to Phase Three at all stages of the justice process (diversion to post-sentence); however, there are restrictions on which statutory bodies can refer at each criminal justice stage. The program is victim-centred, participation is voluntary, all parties must provide informed consent and perpetrators are required to accept some level of responsibility for the offence to be eligible.

# Methods

## Data

This chapter summarises the qualitative analysis of surveys and interviews with 13 victim-survivors who completed an RJ program for a DFV or sexual violence matter. The characteristics of respondents who completed a post-conference survey and/or participated in an interview are summarised in Table 1 below.

### *Post-conference survey*

All 34 victim-survivors who participated in a Phase Three conference between November 2018 and August 2022 were invited to complete a telephone survey approximately six weeks later. Ten respondents (29.4%) completed the survey, including two victim-survivors of sexual violence and eight victim-survivors referred for DFV offences (*n*=2 intimate partner violence; *n*=3 child-to-parent violence; *n*=2 adolescent family violence against an older family member; *n*=1 adolescent family violence against a sibling).

The survey included questions about participants' experiences and satisfaction with the RJ process before, during and after the conference, and any agreements made during the conference. Participants were also asked about their feelings around safety and support, being listened to and respected, and whether they were able to say what they wanted to say. Participants who completed the post-conference survey were also provided with the opportunity to consent to being contacted by researchers to take part in a follow-up interview at a later date.

*Interviews*

Victim-survivors were recruited for an interview in two ways. Either they provided consent to be re-contacted for an interview through the post-conference survey (see above), or they could consent to their contact details being shared with the research team through the Convenor assigned to their matter.

Seven interviews were conducted with victim-survivors of DFV and sexual violence, including two victim-survivors of sexual violence and five victim-survivors of family violence perpetrated by young people. Interviews were conducted in person or via telecommunication, at least three months after they had exited the program. Four interviewees also completed the post-conference survey. Participants were reimbursed $50 for their time.

Participants were asked about their expectations of RJ, their motivations for participating, and how their experience differed from other traditional criminal justice system responses (eg going to court). Participants were also asked about any benefits they received from participating in RJ, and any areas of improvement for the program.

**Table 1: Victim-survivor characteristics (*n*=13 respondents)**

| ID | Gender | Offence type | Year referred | Referrer | Conference | Response provided |
|---|---|---|---|---|---|---|
| CPV1 | F | Child-to-parent violence | 2019 | ACTP | Face-to-face | Survey & interview |
| CPV2 | F | Child-to-parent violence | 2020 | ACTP | Face-to-face | Survey & interview |
| CPV3 | F | Child-to-parent violence | 2019 | ACTP | Face-to-face | Interview |
| CPV4 | F | Child-to-parent violence | 2020 | CYPS | Face-to-face | Interview |
| CPV5 | F | Child-to-parent violence | 2021 | CHC | Face-to-face | Survey |
| FV1 | M | Adolescent violence against older family member | 2019 | ACTP | Video | Survey |
| FV2 | F | Adolescent violence against older family member | 2019 | ACTP | Video | Survey |
| FV3 | F | Adolescent violence against sibling | 2020 | ACTP | Face-to-face | Survey |
| FV4 | F | Adolescent violence against older family member | 2018 | ACTP | Face-to-face | Interview |
| IPV1 | F | Adult violence against current partner | 2019 | MCC | Face-to-face | Survey |
| IPV2 | F | Adult violence against former partner | 2019 | ACTP | Face-to-face | Survey |
| SV1 | F | Sexual violence against adult | 2020 | ACTP | Face-to-face | Survey & interview |
| SV2 | M | Sexual violence against adolescent | 2022 | ACTP | Face-to-face | Survey & interview |

Note: ACTP=ACT Police; CHC=Children's Court; CYPS=Child Youth Protective Services; MCC=Magistrates Court; CPV=child-to-parent violence; FV=family violence; IPV=intimate partner violence; SV=sexual violence

## Analysis

Interview transcripts and the post-conference surveys were analysed in MAXQDA. As this analysis was secondary to the original aims guiding data collection, the guidelines provided by Ruggiano and Perry (2019) for methodological rigour in secondary qualitative data analysis were followed. These guidelines highlight the importance of clarifying the difference between the aims of the parent study (see Lawler, Boxall and Dowling 2025) and those of the secondary analysis.

The parent study was a commissioned evaluation that examined victim-survivors' experiences in each stage of the process (referral, preparation, conference, post-conference) and focused on outcomes and areas for improvement for Phase Three. In this secondary analysis, the aim of examining specific best-practice principles outlined in the literature is narrower than the objectives of the broader evaluation. While there is overlap between the themes and findings across both pieces of work, this study is a distinct piece of research that is less targeted in its aims and conclusions.

The generic inductive approach was applied. This qualitative analytic approach involves reading and re-reading text and coding information based on a combination of predefined categories of interest and other themes that emerge from the data (Thomas 2006). This included recording themes relating to the best-practice principles outlined in the literature (Jeffries, Wood & Russell 2021; Wolthuis & Lünnemann 2016), as well as those that were naturally reflected on and reported by victim-survivors.

## Limitations

Only victim-survivors whose matters proceeded to conference completed the post-conference survey and only 14.8 percent of matters referred to Phase Three proceeded to a conference in the evaluation period ($n$=24, 2018–2022; Lawler, Boxall and Dowling 2025). This means that only a small pool of victim-survivors was eligible to be surveyed and, following on from this, that only a small sample did ($n$=13). The views of those who completed the survey and participated in an interview may not, therefore, be representative of all victim-survivors who participated in Phase Three or of all participants in RJ programs for DFV and sexual violence more broadly.

Different forms of DFV and sexual violence often have distinct underlying dynamics as well as shared associated drivers and impacts. The offence types included in this research represent those eligible for referral into Phase Three, rather than implying that these offences are interchangeable. There were no interviews with victim-survivors who proceeded to conference for intimate partner violence (IPV) or adult family violence (FV) matters. As a result, understanding of the effectiveness of the Scheme for IPV and FV matters is limited, particularly with regard to non-heterosexual relationships.

The COVID-19 pandemic and associated lockdowns across the research period also restricted the number of referrals being made to Phase Three as well as the ability of the Restorative Justice Unit and Convenors to make progress with matters.

This study involves secondary analysis of data collected for the purposes of evaluating the processes and outcomes of an RJ program for DFV and sexual violence. Victim-survivors were asked about their experiences of the service they received, but there may be some aspects of best practice (eg the extent to which Convenors and services worked together) where the participants' depth of understanding may be limited. Therefore, there may be gaps in the information provided by participants.

# Results

Best practices discussed by victim-survivors are categorised into four main themes: Convenor skills and expertise, victim-centred processes, intensive preparation and risk management, and collaboration with other services and experts.

## Convenor skills and expertise

Victim-survivors consistently discussed Convenors' expertise and skills. All victim-survivors reflected positively on the experience of working closely with the Convenors to prepare for the conference, particularly regarding the high level of support they received:

> I guess I just felt supported by them in the process. (FV4, 2018)
>
> Definitely no concerns, surprised how well it went … Good support in lead up was a factor. (IPV2, 2019)

Several victim-survivors said they appreciated Convenors being available and flexible, as well as listening and contacting them regularly to see how they were, for example:

> Well, she would check in with me. And if I needed to talk to her, she was available … she actually listened … she knew how to get me through that situation. (CPV4, 2020)

Multiple victim-survivors said that they found it helpful when Convenors expressed hope and optimism about the potential for change, as they were often experiencing hopelessness about their circumstances. This was particularly the case with child-to-parent violence (CPV) matters:

> I was getting close to losing hope that I will ever have my son back … from the first day [the Convenor] came, hope started again. (CPV2, 2020)
>
> [the Convenors] were exceptional, inspired a lot of confidence. (FV2, 2019)

Several victim-survivors reflected on Convenors' skills in managing perpetrator behaviour:

> There was a couple of awkward times through the process where my son would get a bit frustrated with the whole thing, and he'd want to walk out and things like that … The facilitators were great, and they didn't pressure him to stay but they helped to talk through what was going on for him. They'd even talk outside the room, they'd go out with him and talk to him, and then they'd come back in. They really facilitated the whole thing, I think, really well. (FV4, 2018)

One parent of a young person who had used violence was very positive about the approach taken by the Convenor in supporting the offender to change:

> [the Convenor was able to] make him see that he could be in control without being a bully … [the Convenor] managed to create a relationship with him, and [the offender] genuinely liked her at the end … those people working in restorative justice are more efficient than everything else I tried together. (CPV2, 2020)

Feedback provided by victim-survivors show that it is important to them that Convenors are able to recognise and manage perpetrators' attempts to manipulate the process and are able hold them accountable in real time:

> I believe [the offender] should have been pulled up when [they] started speaking during my opportunity. (IPV1, 2019)
>
> It gave me an opportunity to say how it made me feel in a way that he wasn't going to interrupt me and he had to listen. Whereas, in the home setting if he didn't want to hear it, he would just walk away or distract himself or me with something else. (CPV3, 2019)

Several participants reflected on best practice in the way Convenors managed their expectations. This was evident in descriptions from victim-survivors about how Convenors clearly explained what to expect and provided them with information about the process. In one case, the victim-survivor spoke about how the Convenor was clear from the beginning that the process was going to take time:

> [the Convenor] said, "I would get him there, but you're going to have to be patient and it's going to be little steps," and eventually that's what happened, but little step by little step. (CPV2, 2020)

This was also evident in victim-survivors' reflections on the alignment between their justice needs and the offender's capacity and willingness to meet these:

> I guess [the Convenor's] role was explaining the purpose and the expected outcome, but also understanding how our family worked and what was realistic to [the offender]. (FV4, 2018)

In summary, all victim-survivors reflected positively on Convenors' judgement, skills and expertise, which was apparent in the tailored support they received, and Convenors' ability to manage their expectations about the process. Victim-survivors also valued Convenors' ability to manage problematic behaviours of perpetrators.

## Victim-centred processes

Victim-survivors' experiences of best practice in relation to a victim-centred approach were evident in reports of participatory decision making, a flexible approach to preparation and facilitation that took account of their justice needs and preferences, and a diverse range of desired outcomes being met through the process. This included being given the opportunity to voice the impacts of the crime, ask questions, re-establish feelings of safety, repair relationships, and achieve accountability and forgiveness. Several interviewees observed that everyone involved in the conference benefited from 'being able to say their piece' (CPV1, 2019).

> [It was an] opportunity to talk to him face to face, lift myself up. Did wonders for me and my confidence, part of healing process. (IPV2, 2019)

> I felt heard, even though I didn't need my perpetrator to understand what I was saying. I just needed a space to say it. And I don't think I realised how helpful that was going to be to me. (SV1, 2020)

Multiple victim-survivors reported that the experience of voicing the impacts of the offence and the circumstances surrounding it had flow-on effects for other family members and relationships:

> They helped the family more broadly than just between me and my grandson … The longer lasting outcomes for that is that it had a positive impact on the trajectory of our relationships. It put us in a place where we could keep growing in our relationships, rather than being stalled and estranged from one another. (FV4, 2018)

Two interviewees described the conference as important for building feelings of safety and developing practical safety plans to re-establish contact as well as boundaries:

> The conference I think that was more about getting [the offender] coming back, and working her way back to live with me … to work out how she was going to come back to live with me. Making sure that [the offender] was safe enough to be around, basically. (CPV4, 2020)

Others expressed satisfaction at being able to use the process as a way for the offender to take accountability and apologise, which was important to them. Relatedly, expressing forgiveness was important for some victim-survivors, particularly if the offender was a young person in their care:

> At the conference, it took over a year, but [the offender] said, "I'm sorry," and he meant it. (CPV2, 2020)

> Because I wanted to forgive him. I live in a world where forgiveness is the really powerful thing … I think it was more for his benefit that he apologised so that then, when I said I accepted his apology, he could feel forgiven. (FV4, 2018)

Victim-survivors reported that Convenors' engagement with them was tailored to their needs, consistent with victim-centred and trauma-informed restorative practice. One victim-survivor positively described Convenors' responsiveness to their request for some affected persons to join later in the conference as they did not want to voice certain issues with them present (SV2, 2022). While victim-survivors valued intensive one-on-one support from Convenors in the lead-up to the conference, some reported getting to a point where they preferred less frequent contact:

> Every time you talk about it, you're rehashing it and you're re-experiencing trauma. It is helpful but at a certain point it holds you back … It would be a constant topic in my therapy sessions because it would be a constant thing in my life. In the end, [the Convenors] were lovely. I decided that – we decided together that I didn't need any more meetings until one more meeting before the conference. (SV1, 2020)

Taken together, the importance that victim-survivors' ascribed to the victim-centred nature of the process was evident in their satisfaction with RJ providing them the opportunity to achieve individualised justice in a wide range of circumstances and cases. Victim-survivors particularly highlighted the importance of flexibility and adaptability during preparation and facilitation.

## Intensive preparation and risk assessment

The majority of victim-survivors reported feeling satisfied with certain elements of the conference preparation process, including measures to help them feel ready to participate and help them to understand their role in the process. This was related to Convenor characteristics, specifically support, availability, specialist expertise and understanding (see section above on Convenor skills and expertise):

> [the Convenors] did a lot of preparation. Extraordinary important part of process, big investment, time and support. (FV2, 2019)

Several participants reflected on how high-risk matters were managed well when they described re-contact with the offender being approached slowly and in highly controlled circumstances:

> She'd [the offender] come over for a few hours or something, and then she'd go again. And then she'd like, just gradually coming back slowly, instead of coming back straightaway … working back in, the small steps, yes, that was so much better. (CPV4, 2020)

> The first six months when I was with [the Convenor], he [the offender] did not come in once, and then we had many, many meetings and eventually he came for one night, and a week later two nights, and he started to come back a little bit here and there. (CPV2, 2020)

One victim-survivor reported the Convenors 'did everything they could' to prepare them for the conference, describing the space where they eventually met the offender face-to-face as 'heavily curated' (SV2, 2022). They appreciated how Convenors explained when, where and how the conference would take place, visiting the space with them beforehand, and supporting them to write their talking points on cards in case they needed conversational prompts on the day.

One victim-survivor spoke about how delays were influenced by different people involved in the matter being at varying stages of readiness to meet with each other in a constructive way. However, with ongoing preparation and practice over the course of a year, they described the final conference as 'a beautiful success' (CPV2, 2020). Similarly, another victim-survivor said they appreciated the additional time to prepare:

> [it was] quite a big long process, because there is a lot of stuff to do … it actually took quite a long time. But throughout the time and the process I was actually quite relaxed because it was giving me power. It's a slow and steady process but it's absolutely worth it. (SV2, 2022)

Victim-survivors valued the intensive preparation period as this supported them to feel ready to engage with RJ in their own way. Risk assessment processes such as the careful management and curation of contact between victim-survivors and perpetrators enhanced feelings of safety and support for victim-survivors, illustrating the lived experience of this best practice principle.

## Close collaboration with services and experts

Broadly, the feedback of victim-survivors points to the value they saw in therapeutic support for themselves and perpetrators, both during and after engagement in the RJ process. One victim-survivor spoke about the benefits of engaging in therapy facilitated through RJ:

> The biggest thing that they provided me was a psychologist who actually did cognitive therapy with me, and it was the best thing, it was great. It's made a huge difference in my ability to move on. (FV4, 2018)

Another victim-survivor spoke about how important it was for them to see the person who harmed them engaging in therapy. This helped them to move forward from what happened and empowered them to feel like they had done something to prevent the offender harming others:

> I did genuinely want him to get help before he hurt someone else like how he hurt me. Because I wouldn't wish what he did to me on anyone else … I did get what I wanted from him which was answers and for him to go seek help. (SV2, 2022)

In one matter, a victim-survivor of sexual violence was disappointed the offender had not engaged with any consent education prior to the conference:

> He [the offender] has no understanding of consent. Twelve months between offence and RJ. Thought RJ would have educated him and encourage to do course before RJ. (SV1, 2020)

Other participants reported appreciating how RJ overlapped with other services to provide continuity of support after the formal conference part of the program had been finalised. One victim-survivor spoke about how her child (the offender) had been engaging with a local violence prevention service for men, which had been beneficial:

> I like the idea of [the service] because they've got that complexity at the centre that deals with the mental health as well as the domestic violence for adolescent men. (CPV3, 2019)

Another victim-survivor said they appreciated the ease of the transition from engaging primarily with the Restorative Justice Unit during preparation for the conference to a more intensive engagement with family group therapy:

> So restorative justice was stepping back, and then yeah, family group therapy was stepping in. (CPV4, 2020)

The value victim-survivors saw in therapeutic support is also evident in the criticisms several made about the unavailability of some services. One victim-survivor said the young person who harmed her did not participate in any specialised programs for adolescent girls using FV behaviours because 'there were none' (CPV1, 2019).

In fact, several victim-survivors reported feeling frustrated that they could not access the appropriate and relevant specialist services despite being (at least partly) motivated to participate in RJ for this reason:

> I felt like there is definitely gaps in services. People talk to you for a long time and go, "Oh, I was just going to suggest this, or PCYC or…" At the end of it I go, "I've already tried those things." And they go, "Oh, well, thanks for sharing. See you." (CPV3, 2019)
>
> Everyone was telling me, "We can't help you". (CPV2, 2020)
>
> RJ people very supportive. Very useful but no professional[s] apart from them. (IPV2, 2019)

Overall, the availability and accessibility of professional supports alongside the RJ process was important for victim-survivors, in terms of both their own experience of healing and the level of engagement demonstrated by the offender. While several reflected on the limited availability of specialised programs, the supports that victim-survivors found particularly useful (aside from Convenors) were one-on-one clinical psychological support, family group therapy and local violence prevention programs for young men.

## Discussion

This study examined victim-survivors' views and experiences of an RJ program for DFV and sexual violence, drawing out how these align with best practice principles for delivering RJ in this context. Several widely recognised best practice principles emerged as directly relevant to victim-survivors' experience of RJ, including Convenor skills and expertise, victim-centred processes, intensive preparation, and close and ongoing collaboration with external professional supports.

From the perspective of victim-survivors participating in RJ for DFV and sexual violence, the role of the Convenor 'cannot be overstated' (Bolitho 2015: 273), and their skills and expertise in several regards are critical. This is highlighted by the importance victim-survivors ascribe to Convenors understanding their needs and being available and able to offer tailored support. This aligns with the views of stakeholders that best practice relating to RJ in the context of DFV and sexual violence requires Convenors to have specialist knowledge of and experience in the dynamics and impacts of these crimes (Burns & Sinko 2023; Jeffries, Wood & Russell 2021; Wolthuis & Lünnemann 2016). Evidence of this was present in participants' reflections on the extent to which Convenors were able to recognise and manage problematic behaviours by perpetrators (or not) during the RJ process, and lay the groundwork for longer-term behaviour change. As highlighted, this is an especially acute concern in the context of IPV, which regularly occurs in relationships characterised by persistent manipulation and coercion (Gavrielides 2015; Pennell et al. 2021).

Close collaboration with DFV and sexual violence specialists and other relevant professionals is a critical component of best practice in RJ for complex harm matters (Bolitho 2015; Gavrielides 2015; Jeffries, Wood & Russell 2021). Victim-survivors in this study commonly reported that access to professional support alongside RJ was important to them, both for themselves and the person who offended against them. Valuations of the specialised support provided by Convenors to victim-survivors alongside reflections on the boundaries of their expertise highlights the importance of partnerships between RJ and external specialist support services, particularly specialised therapeutic programs for men using violence. While Convenors typically have clinical training and skills, several experts have cautioned that RJ is not, and should not try to be, a therapeutic program (Daly & Wade 2017; Lopez & Koss 2017). These findings highlight that it can be challenging to demarcate the roles and responsibilities of RJ Convenors versus collaborating external specialists, with overlapping skills, interests and expertise, particularly in regard to victim justice interests such as behaviour change and relationship repair.

RJ for DFV and sexual violence is necessarily adaptive, reflecting the diversity among victim-survivors in what they want to achieve through RJ. Indeed, variation in the extent to which victim-survivors wanted additional support outside of RJ, and the forms of support they preferred, underscores the necessity of other best practice principles such as flexibility and responsiveness to individual needs. This flexibility characteristic of RJ facilitates procedural justice for victim-survivors as they can 'steer' the response with tailored support and without infringing on the rights of the perpetrator (Van Camp & Wemmers 2013). Across diverse cases, personal circumstances, levels of perpetrator accountability and timelines, flexibility and participatory decision-making during preparation and facilitation was central to victim-survivors experiencing RJ as safe, positive and healing.

Lastly, practitioners consistently advocate for intensive risk assessment as essential for best practice when providing RJ in the context of DFV cases (Gavrielides 2015; Wolthuis & Lünnemann 2016). Victim-survivors in this study reflected positively on risk assessment practices, such as Convenors facilitating careful and staged re-contact with the offender. Victim-survivors also benefited from Convenors going over with them everything that would happen in the conference during the preparation period. This is consistent with evidence provided by victim-survivors participating in RJ for DFV in the United Kingdom, who similarly describe feeling empowered through preparatory exercises such as attending mock conferences with Convenors to practise voicing the impacts and to assess potential triggers (Gavrielides 2015).

Overall, victim-survivors report many experiences that align with the best practice principles outlined in the literature and that are specific to RJ in the context of DFV and sexual violence. The findings highlight the central role of Convenors in influencing victim-survivors' experiences of best practice, in line with previous research. Variability in the need for, and availability of, specialist therapeutic support alongside RJ highlights the particular importance of other best practice principles such as flexibility in preparation and approach to facilitate individualised justice for victim-survivors of DFV and sexual violence.

## References

*URLs correct as at March 2025*

Bolitho J 2015. Putting justice needs first: A case study of best practice in restorative justice. *Restorative Justice* 3(2): 256–281. https://doi.org/10.1080/20504721.2015.1069531

Burns CJ & Sinko L 2021. Restorative justice for survivors of sexual violence experienced in adulthood: A scoping review. *Trauma, Violence, & Abuse* 24(2): 340-354. https://doi.org/10.1177/15248380211029408

Daly K 2006. Restorative justice and sexual assault: An archival study of court and conference cases. *British Journal of Criminology* 46(2): 334–356. https://doi.org/10.1093/bjc/azi071

Daly K & Wade D 2017. Sibling sexual violence and victims' justice interests: A comparison of youth conferencing and judicial sentencing. In E Zinsstag & M Keenan (eds), *Restorative responses to sexual violence*. Routledge: 143–178

Gavrielides T 2015. Is restorative justice appropriate for domestic violence cases? *Social Work Review* 14(4): 105–121. https://www.researchgate.net/publication/385337812_Is_Restorative_Justice_Appropriate_for_Domestic_Violence_Cases

Herman JL 2005. Justice from the victim's perspective. *Violence Against Women* 11(5): 571–602. https://doi.org/10.1177/1077801205274450

Jeffries S, Wood WR & Russell T 2021. Adult restorative justice and gendered violence: Practitioner and service provider viewpoints from Queensland, Australia. *Laws* 10(1): 13. https://doi.org/10.3390/laws10010013

Joudo Larsen J 2014. *Restorative justice in the Australian criminal justice system.* Research and public policy series no. 127. Canberra: Australian Institute of Criminology. https://www.aic.gov.au/publications/rpp/rpp127

Jülich S 2006. Views of justice among survivors of historical child sexual abuse: Implications for restorative justice in New Zealand. *Theoretical Criminology* 10(1): 125–138. https://doi.org/10.1177/1362480606059988

Lawler S, Boxall H & Dowling C 2025. *Restorative justice conferencing for domestic and family violence and sexual violence: Evaluation of Phase Three of the ACT Restorative Justice Scheme*. Research Report no. 33. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/rr77680

Lopez EC & Koss MP 2017. The RESTORE Program for sex crimes: Differentiating therapeutic jurisprudence from restorative justice with therapeutic components. In E Zinsstag & M Keenan (eds), *Restorative responses to sexual violence*. Routledge: 212–228

Marsh F & Wager NM 2015. Restorative justice in cases of sexual violence: Exploring the views of the public and survivors. *Probation Journal* 62(4): 336–356. https://doi.org/10.1177/0264550515619571

Marshall T 1999. *Restorative justice: An overview.* London: Home Office

Nascimento AM, Andrade J & De Castro Rodrigues A 2023. The psychological impact of restorative justice practices on victims of crimes: A systematic review. *Trauma, Violence, & Abuse* 24(3): 1929–1947. https://doi.org/10.1177/15248380221082085

Pennell J, Burford G, Sasson E, Packer H & Smith EL 2021. Family and community approaches to intimate partner violence: Restorative programs in the United States. *Violence Against Women* 27(10): 1608–1629. https://doi.org/10.1177/1077801220945030

Robinson G & Shapland J 2008. Reducing recidivism: A task for restorative justice? *British Journal of Criminology* 48(3): 337–358. https://doi.org/10.1093/bjc/azn002

Ruggiano N & Perry TE 2019. Conducting secondary analysis of qualitative data: Should we, can we, and how? *Qualitative Social Work* 18(1): 81. https://doi.org/10.1177/1473325017700701

Sardina A & Ackerman AR 2022. Restorative justice in cases of sexual harm. *CUNY Law Review* 25: 1. https://academicworks.cuny.edu/clr/vol25/iss1/10/

Senate Legal and Constitutional Affairs References Committee 2023. *Current and proposed sexual consent laws in Australia*. https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Legal_and_Constitutional_Affairs/sexualcontentlaws/Report

Silva S 2020. *Impact of victim offender dialogue on victims of serious crimes: A longitudinal cohort-control study.* Center for Victim Research Repository. https://www.ojp.gov/library/publications/impact-victim-offender-dialogue-victims-serious-crimes-longitudinal-cohort

Stubbs J 2007. Beyond apology? Domestic violence and critical questions for restorative justice. *Criminology & Criminal Justice* 7(2): 169–187. https://doi.org/10.1177/1748895807075570

Thomas DR 2006. A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation* 27(2): 237–246. https://doi.org/10.1177/1098214005283748

Van Camp T & Wemmers J-A 2013. Victim satisfaction with restorative justice: More than simply procedural justice. *International Review of Victimology* 19(2): 117–143. https://doi.org/10.1177/0269758012472764

Wolthuis A & Lünnemann K 2016. *Restorative justice and domestic violence: A guide for practitioners.* Saatavilla: Verwey-Jonker Institute. https://www.unodc.org/e4j/data/_university_uni_/restorative_justice_and_domestic_violence_a_guide_for_practitioners.html

This work was originally published as *Trends & issues in crime and criminal justice* no. 720. It is available online at **https://doi.org/10.52922/ti78014**.

# 14. Ransomware targeting individuals and small businesses: Vulnerabilities and impacts

Isabella Voce and Anthony Morgan

Ransomware refers to a type of malicious software that encrypts or blocks access to files until a user has paid a ransom (Australian Cyber Security Centre (ACSC) 2023a). Ransomware is a multi-billion-dollar global industry (Wall 2021), with attacks targeting major infrastructure, government services, businesses and individuals increasing significantly in the past decade. This increase can be attributed to many factors, including the growing reliance of individuals and organisations on their IT and data systems, which make them valuable targets to criminals (Connolly & Wall 2019), and the professional ecosystem that has grown to enable and support ransomware offenders (Wall 2021). This includes easily accessible ransomware toolkits and ransomware-as-a-service, which allow cybercriminals to launch attacks without having advanced technical knowledge themselves (Sharmeen et al. 2020).

Within this changing and growing ransomware landscape, it is crucial to understand how ransomware targets individuals and organisations, including the techniques used and the vulnerabilities exploited. Ransomware research has ballooned in the past decade, yet most of this research has focused primarily on its technical aspects, with little attention to the human factor in victimisation (Connelly & Wall 2019). In particular, few studies have examined the ransomware experiences of small to medium enterprise (SME) owners, who are at high risk of cybercrime victimisation and often have low levels of cybersecurity understanding and practices (Matthijsse et al. 2025), despite most rating cybersecurity as important (ACSC 2023b). Understanding this helps in developing effective prevention strategies and support for victims.

In the 2023 Australian Cybercrime Survey, comprising 13,887 Australian adults, nearly five percent of survey respondents had received a ransom message on their device in the preceding 12 months (Voce & Morgan 2023a). This was an increase on the estimated 2.1 percent of respondents who had received a ransom message on their device in a 2021 Australian survey (Voce & Morgan 2021). Additional analysis of data from the 2023 Australian Cybercrime Survey revealed that SME owners were more likely to have been the victim of ransomware resulting in an encrypted or disrupted device (6.2%) than SME employees (3.2%) and other respondents (1.5%).

Australian SMEs face increasingly sophisticated ransomware attacks targeting their money, data and reputation (ACSC 2023b). While some ransomware groups specialise in 'big game hunting', or attacks against large companies, the majority tend to target SMEs, which have less sophisticated cybersecurity yet enough revenue, data and access to other victims within their supply chain to be valuable (Wall 2021). A study of 5,600 IT professionals in medium to large organisations across 31 countries found that ransomware attacks were on the rise and their complexity and impact were increasing (Sophos 2022). From 2020 to 2021, the proportion of organisations reporting that one or more devices had been impacted by an attack in the previous year increased from 37 percent to 66 percent, and the proportion of attackers successfully encrypting data increased from 54 percent to 65 percent.

Ransomware can be spread through many different means, including phishing or spam emails, password cracking or stolen credentials, remote desktop connection, infecting legitimate websites with malicious elements like links, worms that spread malware across different devices, and ransomware exploit toolkits (Beaman et al. 2021). In Australia, malicious actors persistently scan for any networks with unpatched systems, at times seeking to use these as entry points leading to higher value targets (ACSC 2023a). In 2021–22, the majority of significant incidents ACSC responded to were due to inadequate patching (ACSC 2023a). Different types of ransomware have different impacts on systems, files and devices. Some methods can often by easily resolved (Beaman et al. 2021), while other methods are irreversible (Gómez-Hernández, Álvarez-González & García-Teodoro 2018). Others rely on manipulation and social engineering to exploit the fear of computer users, rather than locking the device or encrypting any data, and install malware that can have longer-term effects (Andronio, Zanero & Maggi 2015).

Cybercriminals are also increasingly targeting the reputation of organisations by threatening to release or sell their data. One-quarter of the ransomware reports made to the ACSC involved 'double extortion', which is when ransomware victims who experience impacts on their devices or systems are also extorted for payment to prevent the data being leaked or sold online (ACSC 2023a). This was evident in 2021 and 2022, when ransomware groups stole and released the personal information of hundreds of thousands of Australians (ACSC 2023a). Such capabilities have been in development since at least 2015, with analysis of 57 variants of ransomware indicating that cybercriminals were increasingly building information-stealing capabilities into their ransomware (Kharraz et al. 2015). Moreover, cybercrime groups have also begun 'naming and shaming' victims on public forums to increase pressure on the victim to make the ransom payment (Wall 2021). This is particularly a risk for SMEs because, aside from any impacts on business systems and devices, breaches of customer data cause significant reputational damage (Biddle, Gray & McEachern 2022).

Traditionally cybercriminals use fully automated mass distribution ransomware, which is very fast in execution and hits as many targets as possible (Brewer 2016). This method is changing, with ransomware attacks becoming more targeted at specific businesses and companies, where there is the potential for higher profits (Europol 2018). Targeted ransomware attacks involve more sophisticated attackers who spend more time investigating a targeted system's design, cause significant system and device damage and disruption, and demand larger payments from victims (Al-Hawawreh, Den Hartog & Sitnikova 2019), causing more damage than opportunistic attacks (Connolly et al. 2020). The types of systems, devices and data that can be impacted by ransomware are also becoming more diverse (Al-Hawawreh, Den Hartog & Sitnikova 2019).

We know there are certain risk factors for ransomware victimisation. Data breaches have been closely linked to ransomware attacks (Europol 2021). Analysis of 2021 survey data from the Australian public found that respondents who had been notified of a data breach were 79.5 percent more likely to have also received a ransom message on their device (Morgan & Voce 2022). Further, being the victim of ransomware once increases the likelihood of becoming a repeat victim (Barracuda 2023). In particular, paying ransoms can incentivise attackers to target victims multiple times (Department of Home Affairs 2021).

Despite this evidence, our understanding of what makes individuals and organisations fall victim to ransomware is not well developed (Holt & Bossler 2014). There is a lack of studies examining the experiences of businesses and individuals that have fallen victim to ransomware attacks (Connolly et al. 2020). We aim to address this gap by using data from a sample of 331 ransomware victims who participated in the 2023 Australian Cybercrime Survey to understand the events leading up to the incident, the vulnerabilities exploited by perpetrators and the impacts on devices, systems and data for both work-related incidents and personal incidents.

# Method

## Australian Cybercrime Survey

The sample for this study was drawn from the 2023 Australian Cybercrime Survey, which was a national survey of 13,887 online Australians aged 18 years and over that measured cybercrime trends, victim characteristics, help-seeking behaviour, financial losses and other harms. Questions about ransomware victimisation were included in a section of the survey focused on malware. The survey was conducted by Roy Morgan Research in early 2023 using its Single Source panel and panels managed by PureProfile and Dynata. Proportional quota sampling, a non-probability sampling method, was used to ensure the sample was demographically representative of the Australian population. Quotas were based on the Australian adult population stratified by age, sex and usual place of residence, derived from Australian Bureau of Statistics population data. The data were subsequently weighted by age, sex and usual place of residence to ensure the data were representative of the spread of the Australian population. Additional random iterative method weights were applied to correct for education level, internet use and social media use. All findings presented in this report are based on weighted data.

Respondents who said they had received a ransom message on their device in the 12 months prior to the survey were directed to an addendum with follow-up questions. Respondents were asked about the most recent incident of ransomware they had experienced, including:

- how they received the ransom message;
- what impacts occurred to their devices, systems and data around the time of receiving the message;
- the amount demanded in the ransom message and the time limit given for payment;
- their reasons for deciding to pay or not; and
- actions taken to resolve, remedy or report the incident.

Because this chapter is focused on the most recent incident reported by victims rather than incidents that occurred over the previous 12 months, results differ from those reported by Voce and Morgan (2023a).

## Sample

In the 12 months prior to the survey, 673 respondents had received a ransom message on their device demanding payment (with or without device disruption or compromise). Of these victims, 331 participants indicated that either of the following had occurred:

- ‘My systems, devices or files had a virus or were inaccessible (eg locked or unreadable) and I received instructions for paying a ransom to restore access’; or
- ‘My devices, servers, service or networks were disrupted (eg slowed down, lost connection, had outages) and I received instructions for paying a ransom to restore functionality.’

Victims who experienced device, system or data disruption or compromise were included in the current analysis. The remaining 342 victims, who said their data had been stolen and they had to pay to prevent the information being sold or leaked online but who did not report that their device had been disrupted or compromised, were excluded from the current study. Of the 331 victims in the current sample:

- 110 respondents (33%) owned an SME, which was defined as a business with fewer than 200 employees;
- 219 respondents (66%) did not own an SME but may have been employed by one; and
- one respondent did not indicate their employment status (and was therefore excluded from any comparative analysis).

## Limitations

To identify the sample for this study, we used a bottom-up approach to asking about victimisation. This was to ensure that identifying as a victim did not rely on a non-technical audience understanding the definition of ransomware. Given that cybercriminal modus operandi are constantly changing, the incident descriptions used in this study may not fully capture every type of ransomware, and some incidents may not easily fit with these categories. Relatedly, we excluded respondents who were the victim of hacking and data extortion where there were no signs indicating malware on their devices or systems, since the focus was on pure ransomware, but we note that the experiences of this other group may be different to those reported by victims whose devices were impacted.

Although respondents were assured that their responses to the survey were anonymous and private, some respondents may nevertheless have been excluded from the current sample because of their reluctance to disclose experiences of victimisation due to shame or embarrassment. This is particularly the case for SME owners or employees who experienced a work-related ransomware attack on their business. We also note the potential limitations of relying on respondents to accurately recall information about an incident that could have occurred up to 12 months prior to the survey.

# Results

## Receiving the ransom message

More than half of all ransomware victims had received more than one ransom message in the 12 months prior to the survey (51.7%, see Table 1), with SME owners significantly more likely to receive multiple ransom messages than other victims (63.7% vs 46.2%; $F(1, 246)=5.35$, $p<0.05$). When asked about the most recent incident, victims most commonly received the ransom message on their mobile phone (56.0%) or laptop or computer (49.1%), and this usually came in the format of an email, text or private online chat (59.1%) or pop-up notification (48.1%). There were no significant differences between SME owners and other victims in terms of which device the ransom appeared on or format of the ransom.

**Table 1: Frequency, target and format of ransom messages**

| | SME owner | | Not an SME owner | | Total | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| **Number of ransom messages received in the 12 months prior to survey** | | | | | | |
| One | 26 | 36.3 | 83 | 53.8 | 109 | 48.3 |
| Two | 18 | 25.9 | 26 | 16.9 | 44 | 19.7 |
| Three | 18 | 25.2 | 17 | 10.8 | 35 | 15.3 |
| Four | 5 | 7.2 | 14 | 9.0 | 19 | 8.4 |
| Five or more | 4 | 5.5 | 15 | 9.6 | 19 | 8.3 |
| **The device on which the victim received message/s in the most recent incident[a]** | | | | | | |
| Mobile phone | 56 | 61.4 | 102 | 53.4 | 158 | 56.0 |
| Laptop or computer | 44 | 48.4 | 94 | 49.4 | 138 | 49.1 |
| Tablet | 13 | 13.8 | 17 | 8.9 | 29 | 10.5 |
| Other | 0 | 0 | 1 | 0.6 | 1 | 0.4 |
| **The format of the message/s in most recent incident[a]** | | | | | | |
| Message in an email, text or private online chat | 62 | 66.6 | 102 | 55.3 | 164 | 59.1 |
| Pop-up notification on device | 43 | 45.8 | 91 | 49.2 | 133 | 48.1 |
| Other | 3 | 3.6 | 6 | 3.2 | 9 | 3.3 |

a: Respondents could select multiple response options

Note: Excludes respondents who did not answer the question

Source: Australian Cybercrime Survey 2023 [weighted data]

## Events precipitating the ransomware incident

Victims were asked whether they remembered clicking on any suspicious links, pop-ups, buttons, files or attachments prior to their device being impacted (Table 2). Only 43.2 percent of victims could specifically remember clicking on any suspicious elements prior to the incident, and while the difference was not statistically significant this was more common among SME owners than other victims (53.8% vs 37.8%; $F(1, 198)=3.71$, $p=0.06$). Victims most commonly clicked on a suspicious link, pop-up, button, file or attachment on a webpage (52.6%) or in an email (42.3%), followed by a text message (33.5%), a social media post (16.0%) or a private chat platform (13.6%). There were no statistically significant differences between SME owners and other victims in where they clicked on suspicious links, pop-ups, buttons, files or attachments.

**Table 2: Events precipitating the most recent ransomware incident**

| | SME owner | | Not an SME owner | | Total | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| **Whether they remember clicking on a suspicious link, pop-up, button, file or attachment[a]** | | | | | | |
| Victim remembered clicking on any suspicious links, pop-ups, buttons, files or attachments | 33 | 53.8 | 28 | 37.4 | 78 | 43.2 |
| **Where they clicked on the suspicious link, pop-up, button, file or attachment[a,b]** | | | | | | |
| On a webpage | 17 | 51.2 | 17 | 59.8 | 40 | 52.6 |
| In an email | 13 | 39.9 | 14 | 50.5 | 32 | 42.3 |
| In a text message | 11 | 33.4 | 10 | 36.0 | 26 | 33.5 |
| On a social media post | 3 | 9.3 | 5 | 19.1 | 12 | 16.0 |
| On a private chat platform | 4 | 11.2 | 3 | 12.3 | 10 | 13.6 |
| Other (please specify) | 3 | 8.6 | 0 | 0 | 3 | 3.7 |
| **Data breaches** | | | | | | |
| Victim had personal data leaked in previous year | 56 | 58.6 | 95 | 48.7 | 151 | 51.9 |
| **Previous ransom payments[c]** | | | | | | |
| Victim paid the ransom in a separate incident in the past year | 15 | 40.5 | 6 | 10.0 | 21 | 21.5 |

a: Limited to respondents who said they experienced loss of device functionality, impacts on data or files, a virus or software installed on devices, or disruption to networks, servers and services ($n=214$)

b: Respondents could select multiple options. Limited to respondents who remembered clicking on a suspicious link, pop-up, button, file or attachment ($n=78$)

c: Limited to respondents who received more than one ransom message in the 12 months prior to survey ($n=117$)

Note: Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

Over half of all ransomware victims reported having had their data exposed in a data breach in the 12 months prior to the survey (51.9%). While this was more common among SME owners (58.6%) than other victims (48.7%), this difference was not statistically significant.

Among the victims who reported receiving ransom messages multiple times in the 12 months prior to survey, 21.5 percent had paid a ransom following one of these previous threats. In comparison to other victims, SME owners were much more likely to have paid the ransom in response to a previous threat (40.5% vs 10.0%; $F(1, 108)=9.37$, $p<0.01$).

## Impact on devices, systems and data

As shown in Table 3, the most common impacts reported by victims were that their servers, services or networks were disrupted (37.1%); data or files on their devices were altered, removed or locked (30.1%); or the device itself lost functionality (29.6%). One in five victims stated that their device had software or a virus installed (21.9%), and a quarter said that they experienced other unspecified impacts (24.0%). There were no statistically significant differences between SME owners and other victims.

**Table 3: Impacts on devices and systems at the time of receiving the ransom message**

| | SME owner | | Not an SME owner | | Total | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| My servers, service or networks were disrupted (eg slowed down, lost connection, had outages) | 39 | 41.0 | 65 | 35.1 | 104 | 37.1 |
| Data or files on the device(s) were altered, removed or locked (eg inaccessible or password protected) | 34 | 35.3 | 51 | 27.4 | 85 | 30.1 |
| The device(s) lost functionality (eg it was locked, inaccessible or password protected) | 28 | 28.9 | 55 | 29.9 | 83 | 29.6 |
| Software or a virus was installed on my device(s) | 20 | 21.0 | 41 | 22.5 | 62 | 21.9 |
| Other | 22 | 23.3 | 45 | 24.4 | 67 | 24.0 |

Note: Respondents could select multiple options. Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

The devices most commonly impacted were the victim's laptop computer (46.3%), smartphone (40.7%), desktop computer (22.4%) and tablet or iPad (22.3%). Compared with other victims, SME owners were more likely to experience impacts to their smart wearables (21.5% vs 6.1%; $F(1, 229)=8.33$, $p<0.01$) and security alarms or intercom entry systems (16.7 vs 1.9%; $F(1, 299)=19.711$, $p<0.001$; Table 4).

**Table 4: Devices or systems that lost functionality or connection, had a virus, or had files or data blocked**

| | SME owner | | Not an SME owner | | Total | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| Laptop computer | 37 | 49.8 | 60 | 44.4 | 97 | 46.3 |
| Smartphone (eg iPhone, Samsung Galaxy) | 26 | 35.4 | 59 | 43.6 | 85 | 40.7 |
| Desktop computer | 20 | 27.5 | 26 | 19.6 | 47 | 22.4 |
| Tablet or iPad | 22 | 30.1 | 24 | 18.1 | 47 | 22.3 |
| Smart TV | 14 | 19.4 | 15 | 10.9 | 29 | 13.9 |
| Smart wearables (eg Fitbit, smart watch)** | 16 | 21.5 | 8 | 6.1 | 24 | 11.5 |
| A gaming console connected to the internet or your TV (eg PlayStation, Xbox) | 13 | 17.9 | 7 | 5.4 | 21 | 9.8 |
| Modem/router | 10 | 13.1 | 7 | 5.5 | 17 | 8.2 |
| Smart home devices (eg Google home, Amazon Home Family) | 9 | 12.5 | 7 | 5.3 | 16 | 7.9 |
| Smart security/alarm/intercom entry system*** | 12 | 16.7 | 2.6 | 1.9 | 15 | 7.2 |
| Servers or networks | 5 | 7.0 | 6 | 4.5 | 11 | 5.4 |
| Home automation (eg appliances, lights etc) | 3 | 3.8 | 2 | 1.8 | 5 | 2.5 |
| Webpages/websites | 2 | 2.7 | 0 | 0 | 2 | 1.0 |
| Other | 0 | 0 | 1 | 0.9 | 1 | 0.6 |

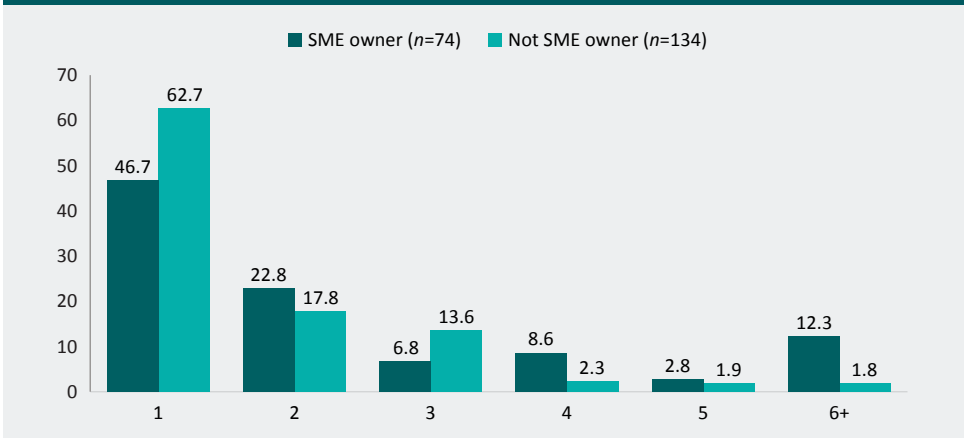
***statistically significant at $p<0.001$, **statistically significant at $p<0.01$

Note: Limited to respondents who said they experienced loss of device functionality, impacts on data or files, a virus or software installed on devices, or disruption to networks, servers and services ($n=214$). Respondents could select multiple options. Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

Overall, SME owners experienced impacts across significantly more types of devices than other victims ($F(4, 1008)=2.90$, $p<0.05$), with 23.7 percent experiencing impacts on four or more devices compared with just six percent of other victims (See Figure 1).

**Figure 1: Number of device types impacted (%)**




Note: Limited to respondents who said they experienced loss of device functionality, impacts on data or files, a virus or software installed on devices, or disruption to networks, servers and services (*n*=214). Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

As shown in Table 5, SME owners were significantly more likely than other victims who were currently working to report that the impacted device had been issued to them by their workplace (40.1% vs 19.5%; $F(1, 163)=6.31$, $p<0.05$) or that they used the personal device for work or business-related tasks (52.6% vs 28.1%; $F(1, 170)=8.38$, $p<0.01$). A higher proportion of SME owners said the data or files impacted by the incident were owned or used by their business or place of employment (49.2% vs 33.1%), but given the relatively small numbers this difference was not statistically significant.

**Table 5: Features of ransomware incidents related to work**

| | SME owner | | Not an SME owner | | Total | |
|---|---|---|---|---|---|---|
| | *n* | % | *n* | % | *n* | % |
| The device impacted was issued by workplace* | 26 | 40.1 | 14 | 19.5 | 41 | 29.2 |
| The impacted personal device was also used for work- or business-related tasks** | 36 | 52.6 | 23 | 28.1 | 59 | 39.1 |
| The data or files impacted by the incident were owned or used by business or place of employment[a] | 14 | 49.2 | 10 | 33.1 | 24 | 40.9 |

**statistically significant at $p<0.01$, *statistically significant at $p<0.05$

a: Limited to people who had data or files on the devices altered, removed or locked (*n*=85)

Note: Limited to respondents who were currently working (*n*=251). Limited to respondents who said they experienced loss of device functionality, impacts on data or files, a virus or software installed on devices, or disruption to networks, servers and services (*n*=214). Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

Victims who indicated that the ransomware event was work-related were asked whether the ransomware had spread to other devices, systems or email accounts in their workplace. SME owners were significantly more likely than other victims to report that the ransomware had spread to other workplace devices, systems or email accounts (52.5% vs 24.0%; $F(1, 86)=6.32$, $p<0.05$).

## Features of the threats

Nearly a third of ransomware victims (29.6%) were threatened that their stolen data would be sold or shared if they did not pay the ransom. This was similar for SME owners (31.7%) and other victims (28.6%).

Ransomware perpetrators demanded payments in money, cryptocurrency and gift cards. The median total amount demanded was $300 for SME owners and $758 for other respondents. The mean value demanded was higher for SME owners ($12,287) than for respondents who were not SME owners ($7,213) for although this difference was not statistically significant (see Table 6).

**Table 6: Ransom amounts and forms of payment demanded**

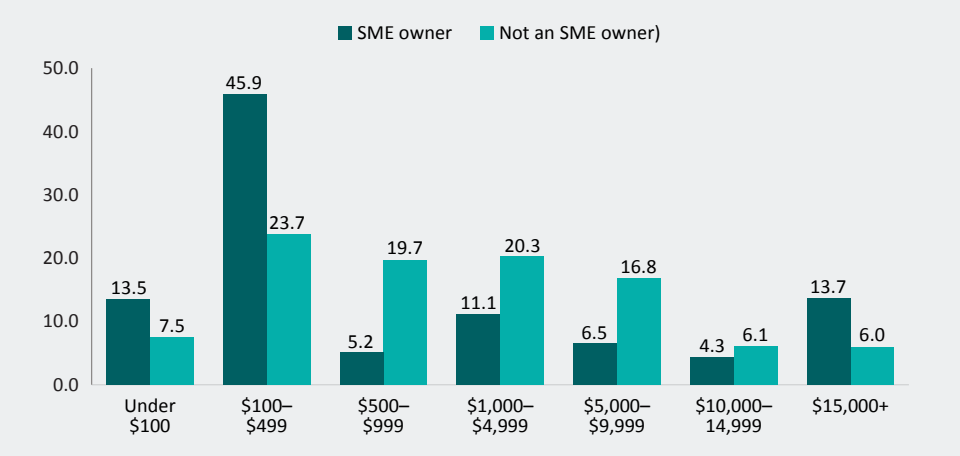
| | | SME owner | Not an SME owner | Total |
|---|---|---|---|---|
| Money | Median (mean) | $100 ($4,723) | $500 ($3,346) | $500 ($3,779) |
| | Range | $2–$50,000 | $3–$100,000 | $2–$100,000 |
| Cryptocurrency | Median (mean) | $100 ($9,845) | $986 ($13,385) | $500 ($11,986) |
| | Range | $5–$500,000 | $5–$350,000 | $5–$500,000 |
| Gift cards | Median (mean) | $100 ($1,714) | $100 ($734) | $100 ($1,184) |
| | Range | $1–$20,000 | $1–$6,999 | $1–$20,000 |
| Total | Median (mean) | $300 ($12,287) | $758 ($7,213) | $555 ($8,728) |
| | Range | $1–$500,000 | $3–$350,000 | $1–$500,000 |

Note: No statistically significant differences were found between SME owners and other respondents. Excludes 169 respondents who did not respond (62 SME owners, 106 respondents who were not SME owners and 1 respondent who did not indicate whether they owned an SME)

Source: Australian Cybercrime Survey 2023 [weighted data]

The amounts demanded of SME owners were relatively small, with 59.4 percent being asked to pay less than $500 (see Figure 2). A small proportion of SME owners received demands for sums of $15,000 or more (13.7%). The amounts demanded of victims who were not SME owners varied greatly, with 80.5 percent of ransoms falling between $100 and $9,999. These differences in the ransom demands received by SME owners and other victims were statistically significant ($F(6, 1005)=3.38$, $p<0.01$).

**Figure 2: Ransom amounts demanded (%)**




Note: Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

Victims were asked whether the ransom message had given them a time limit within which to pay (see Table 7). Half of the ransom messages to SME owners (49.5%) gave a time limit, compared with 39 percent of messages to other victims. Of the 109 victims who were given a time limit for payment, SME owners were given significantly longer timeframes than victims who were not SME owners (median 30 days vs 1 day; $B=0.004$, $F(1, n=102)=14.88$, $p<0.001$).

**Table 7: Time given for ransom payments**

| | | SME owner | Not an SME owner | Total |
|---|---|---|---|---|
| Days | Median (mean) | 30 (52) | 1 (7) | 2 (23) |
| | Range | 0–360 | 0–68 | 0–360 |

Note: Excludes respondents who did not respond

Source: Australian Cybercrime Survey 2023 [weighted data]

Finally, victims were asked whether they had paid or attempted to pay the ransom in the most recent incident. Overall, relatively few respondents indicated that they had actually paid the ransom (12.2%). However, SME owners were more likely than other respondents to have paid the ransom (22.6% vs 7.8%, $F(1, 307)=10.22$, $p<0.01$).

## Discussion

This chapter draws on data from a sample of 331 Australian ransomware victims to better understand how ransomware attacks occur, how victims are targeted and the vulnerabilities exploited by cybercriminals. We had a specific focus on SME owners in light of recent evidence showing they are being targeted more often than large companies (Wall 2021), with attacks that are increasingly frequent, sophisticated and harmful (ACSC 2023b; Sophos 2022).

Typically, ransom messages first appeared on the victim's mobile phone or computer in the form of an email, text message or private online chat. More than half of all ransomware victims had been contacted multiple times over the previous 12 months with threats and demands for payment. This was significantly more common for SME owners, with nearly two-thirds being contacted more than once and more than third contacted three or more times. Importantly, over 40 percent of SME owners had paid in response to one of these previous ransom messages, a significantly higher proportion than among other victims. SME owners were also more likely to have paid following the most recent ransomware incident. This fuels the ransomware business model and can make SME owners appear to be easy to scare and manipulate, increasing their chances of repeat victimisation. Cybercriminals reportedly share between themselves 'sucker lists' of individuals and organisations who have made previous payments (Connolly & Borrion 2022). For this reason, along with the lack of a guarantee that files and systems will be restored and data will not be sold or shared, the Australian Government's advice is to never pay a ransom (ACSC 2023c). Stronger messaging is needed to ensure that SME owners understand and follow this advice.

More than half of all ransomware victims could not remember clicking on any suspicious links, pop-ups, buttons, files or attachments prior to their device being impacted. This speaks to the sophistication of malware attacks, which makes it increasingly hard to spot malicious elements hidden in websites, texts, emails and social media posts. The impacts may not be felt until days or weeks later, meaning attributing the ransomware attack to clicking on something specific can be difficult. For those who could remember, the suspicious element was most commonly in a webpage or email. Ransomware infections through email and webpages are common (Beaman et al. 2021), and prevention requires a combination of technological and human defences. There was some evidence that SME owners were more likely to have clicked on suspicious elements. Business owners are easily identifiable targets as they often have a more visible online presence than other potential victims. Many have public-facing websites, social media pages and email addresses, allowing anyone to contact the business for quotes or enquiries. Preventing ransomware attacks against small business requires cybersecurity measures, but there is an important human element.

Third-party data breaches were a common concurrent event for this sample of ransomware victims. Over half of the sample had experienced a third-party data breach in the 12 months prior to the survey, which is higher than the prevalence across the general Australian Cybercrime Survey sample, of whom 33.6 percent had experienced a data breach (Voce & Morgan 2023a). This is consistent with research showing that third-party data breaches significantly increase the likelihood of ransomware victimisation (Morgan & Voce 2022). A recent survey by the Office of the Australian Information Commissioner (2023) found that 52 percent of victims of third-party data breaches reported an increase in calls and spam texts, which can include hidden malware. Reducing data breaches and putting in place measures to protect affected individuals will help reduce ransomware victimisation.

Business owners experienced far-reaching impacts on their devices, with a higher number of devices and multiple types of devices impacted. These were most often laptop computers and smartphones; however, SME owners were significantly more likely to experience impacts to their smart wearables and security alarms and intercom entry systems—the latter posing clear risks to the physical security of their business. SME owners were also more likely to report that the impacted device had been issued to them by their workplace or that they used a personal device for work- or business-related tasks. They were also more likely to report that the ransomware had spread to other workplace devices, systems or email accounts.

Taken together, these findings reflect the variety of devices and systems that SME owners now use for personal and work-related tasks. With the growing availability of connected devices and systems, SME owners are using this technology to increase collaboration, productivity and revenue (AlphaBeta 2019). Almost all Australian businesses are connected to the internet, more than half have adopted cloud technology and over half have a web presence (Productivity Commission 2023), with one in three SMEs receiving orders via online platforms (AlphaBeta 2019). SMEs are constantly under pressure to be innovative and responsive and are increasingly connected to their customers, employees and supply chain (Baillette & Barlette 2018). They also have more difficulty managing information security because they lack technical and financial resources, and are increasingly targeted by cybercriminals because they can exploit their security weaknesses to access larger organisations of greater interest (Baillette & Barlette 2018).

SMEs and their employees often use their personal devices for work, including personal laptops, tablets and mobile phones. This phenomenon has been termed BYOD (bring your own device), and has become popular in recent years, particularly since working from home became common during the COVID-19 pandemic. While BYOD practices offer flexibility and productivity, they are considered one of the greatest security threats faced by SMEs (Baillette & Barlette 2018). When SME owners moved to working from home during the pandemic, their risk of falling victim to cybercrime increased (Voce & Morgan 2023b). When employees are using their own devices, they may not update the software (Singh 2012) and may be less likely to comply with internal security rules or policies (Hovav & Putri 2016). Guidance on how to manage BYOD-related security issues, especially for those working from home, may help mitigate these risks.

Cybercriminals draw on a number of techniques to extract payment from victims. This includes double extortion, where cybercriminals extort the victim for both device decryption and the non-publication of data. Nearly one-third of all victims were threatened that their stolen data would be sold or shared if they did not pay the ransom. This is higher than previous estimates (closer to one in eight, see Voce & Morgan 2023a), which highlights the importance of asking ransomware victims specifically about the threat of double extortion. Effective information management and data security practices may help better prepare individuals and SME owners for ransomware attacks and ensure they have the confidence to respond appropriately to the threat of double extortion attempts.

Taken as a whole, this study illustrates the vulnerabilities among Australian computer users to the threat of ransomware. Consistent with international research (Matthijsse et al. 2025; Wall 2021), we found that SME owners are viewed as lucrative targets and are especially vulnerable. Efforts to reduce ransomware targeting small businesses (and to limit the potential flow-on consequences for larger organisations that could be infiltrated) need to raise awareness of the ways in which ransomware attacks often happen. This includes education on how to spot suspicious links in emails or on websites, how to respond appropriately to third-party data breaches, and what security measures should be implemented when working from home or using personal devices for work. This also includes supporting victims to identify and remove malware that has spread across devices and systems, and options to mitigate the situation that do not include paying the ransom. This study also makes clear that awareness-raising efforts alone may be insufficient to prevent attacks, meaning there must also be technological solutions that can help protect business owners and avoid the wide-ranging harms that can be inflicted on small businesses (Voce & Morgan 2023a).

## References

*URLs correct as at July 2025*

Al-Hawawreh M, Den Hartog F & Sitnikova E 2019. Targeted ransomware: A new cyber threat to edge system of Brownfield Industrial Internet of Things. *IEEE: Internet of Things Journal* 6(4): 7137–7151. https://doi.org/10.1109/JIOT.2019.2914390

AlphaBeta 2019. Australia's digital opportunity: Growing a $122 billion a year industry. https://digi.org.au/digitalopportunity/

Andronio N, Zanero S & Maggi F 2015. HelDroid: Dissecting and detecting mobile ransomware. In H Bos, F Monrose & G Blanc (eds), *Research in attacks, intrusions, and defenses*. https://www.springerprofessional.de/en/heldroid-dissecting-and-detecting-mobile-ransomware/6878340

Australian Cyber Security Centre (ACSC) 2023a. *ASD cyber threat report 2022–2023*. Canberra: ACSC. https://www.cyber.gov.au/about-us/view-all-content/reports-and-statistics/asd-cyber-threat-report-july-2022-june-2023

Australian Cyber Security Centre (ACSC) 2023b. *Cyber security and Australian small businesses: Results from the Australian Cyber Security Centre Small Business Survey*. Canberra: ACSC. https://www.cyber.gov.au/sites/default/files/2023-03/2023_ACSC_Cyber%20Security%20and%20Australian%20Small%20Businesses%20Survey%20Results_D1.pdf

Australian Cyber Security Centre (ACSC) 2023c. *So, you've been held to ransom?* Canberra: ACSC. https://www.cyber.gov.au/sites/default/files/2023-03/ACSC_Ransomware_Emergency_Response_One_Page_Guide.pdf

Baillette P & Barlette Y 2018. BYOD-related innovations and organizational change for entrepreneurs and their employees in SMEs. *Journal of Organizational Change Management* 31(4): 839–851. https://doi.org/10.1108/JOCM-03-2017-0044

Barracuda 2023. *2023 ransomware insights: The prevalence and impact of ransomware attacks around the world*. https://www.barracuda.com/reports/ransomware-insights-report-2023

Beaman C et al. 2021. Ransomware: Recent advances, analysis, challenges and future research directions. *Computers & Security* 111: 102490. https://doi.org/10.1016/j.cose.2021.102490

Biddle N, Gray M & McEachern S 2022. *Public exposure and responses to data breaches in Australia: October 2022*. ANU Centre for Social Research and Methods. Canberra: Australian National University. https://polis.cass.anu.edu.au/research/publications/public-exposure-and-responses-data-breaches-australia-october-2022

Brewer R 2016. Ransomware attacks: Detection, prevention and cure. *Network Security* 9: 5–9. https://doi.org/10.1016/s1353-4858(16)30086-1

Connolly AY & Borrion H 2022. Reducing ransomware crime: Analysis of victims' payment decisions. *Computers & Security* 119: 102760. https://doi.org/10.1016/j.cose.2022.102760

Connolly LY & Wall DS 2019. The rise of crypto-ransomware in a changing cybercrime landscape: Taxonomising countermeasures. *Computers & Security* 87: 101568. https://doi.org/10.1016/j.cose.2019.101568

Connolly LY, Wall DS, Lang M & Oddson B 2020. An empirical study of ransomware attacks on organizations: An assessment of severity and salient factors affecting vulnerability. *Journal of Cybersecurity* 6(1). https://doi.org/10.1093/cybsec/tyaa023

Department of Home Affairs 2021. *Ransomware Action Plan*. Canberra: Department of Home Affairs. https://www.homeaffairs.gov.au/about-us/our-portfolios/cyber-security/strategy/australias-ransomware-action-plan

Europol 2021. *Internet organised crime threat assessment 2021*. https://www.europol.europa.eu/publications-events/main-reports/internet-organised-crime-threat-assessment-iocta-2021

Europol 2018. *Internet organised crime threat assessment 2018*. https://www.europol.europa.eu/internet-organised-crime-threat-assessment-2018

Gómez-Hernández J, Álvarez-González L & García-Teodoro P 2018. R-Locker: Thwarting ransomware action through a honeyfile-based approach. *Computers & Security* 73: 389–398. https://doi.org/10.1016/j.cose.2017.11.019

Holt TJ & Bossler AM 2014. An assessment of the current state of cybercrime scholarship. *Deviant Behavior* 35(1): 20–40. https://doi.org/10.1080/01639625.2013.822209

Hovav A & Putri FF 2016. This is my device! Why should I follow your rules? Employees' compliance with BYOD security policy. *Pervasive and Mobile Computing* 32: 35–49. https://doi.org/10.1016/j.pmcj.2016.06.007

Kharraz A, Robertson W, Balzarotti D, Bilge L & Kirda E 2015. Cutting the Gordian knot: A look under the hood of ransomware attacks. In M Almgren, V Gulisano & F Maggi (eds), *Detection of intrusions and malware, and vulnerability assessment*: 3–24. https://doi.org/10.1007/978-3-319-20550-2_1

Matthijsse SR, Moneva A, van 't Hoff-de Goede MS & Leukfeldt ER 2025. Examining ransomware payment decision-making among small- and medium-sized enterprises. *European Journal of Criminology* 22(4): 625–645. https://doi.org/10.1177/14773708241285671

Morgan A & Voce I 2022. *Data breaches and cybercrime victimisation*. Statistical Bulletin no. 40. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sb78832

Office of the Australian Information Commissioner 2023. *Notifiable data breaches report: January to June 2023.* Sydney: OAIC. https://www.oaic.gov.au/privacy/notifiable-data-breaches/notifiable-data-breaches-publications/notifiable-data-breaches-report-january-to-june-2023

Productivity Commission 2023. *5-year Productivity Inquiry: Australia's data and digital dividend*. Inquiry report vol 4. Canberra: Productivity Commission. https://www.pc.gov.au/inquiries/completed/productivity/report

Sharmeen S, Ahmed YA, Huda S, Koçer BŞ & Hassan MM 2020. Avoiding future digital extortion through robust protection against ransomware threats using deep learning based adaptive approaches. *IEEE Access* 8: 24522–24534. https://doi.org/10.1109/ACCESS.2020.2970466

Singh N 2012. B.Y.O.D. Genie is out of the bottle – 'devil or angel'. *Journal of Business Management & Social Sciences Research* 1(3): 1–12

Sophos 2022: *The state of ransomware*. https://www.sophos.com/en-us/whitepaper/state-of-ransomware

Voce I & Morgan A 2021. *Ransomware victimisation among Australian computer users*. Statistical Bulletin no. 35. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sb78382

Voce I & Morgan A 2023a. *Cybercrime in Australia 2023*. Statistical Report no. 43. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sr77031

Voce I & Morgan A 2023b. Online behaviour, life stressors and profit-motivated cybercrime victimisation. *Trends & issues in crime and criminal justice* no. 675. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti77062

Wall DS 2021. The transnational cybercrime extortion landscape and the pandemic: Changes in ransomware offender tactics, attack scalability and the organisation of offending. *European Law Enforcement Research Bulletin*, Special Conference edition No. 5. https://ssrn.com/abstract=3908159

This work was originally published as *Trends & issues in crime and criminal justice* no. 724. It is available online at **https://doi.org/10.52922/ti78106**.

# 15. Developing a harm index for individual victims of cybercrime

Isabella Voce and Anthony Morgan

Measuring the scale of harms experienced by cybercrime victims is important in changing attitudes towards the seriousness of cybercrime as a problem and in ensuring the necessary financial, human and technical resources are invested in prevention and responses. Cybercrime harms are often expressed in terms of the financial costs to individuals, businesses and governments. The total cost of cyber-dependent crime to individuals in Australia in 2019 was estimated to be $3.5b (Teunissen, Voce & Smith 2021). The combined losses of scams reported to Scamwatch, ReportCyber and other sources—many of which will have involved an online component—was more than $3.1b in 2022 (Australian Competition and Consumer Commission 2023).

But a growing body of evidence shows victims of cybercrime experience harms that extend beyond these financial losses (Ignatuschtschenko 2021). Cybercrime victims can experience short-and long-term effects such as financial hardship, emotional distress, loss of confidence in other people, physical symptoms (eg difficulty sleeping and poor health), relationship and family conflict and, at the extreme end of the spectrum, suicidal thoughts and ideation (Cross, Richards & Smith 2016; Golladay & Holtfreter 2017; Jansen & Leukfeldt 2018; Whitty & Buchanan 2015). These consequences can vary according to the type of cybercrime, the relationship between the victim and offender, the personal circumstances of the victim prior to falling victim, and the response by authorities (Borwell, Jansen & Stol 2022).

While financial losses are often a more tangible measure of the impact of cybercrime, they also appeal because they allow for different cybercrime types to be compared and measured over time. Developing a metric by which the full range of harms can be measured has several benefits. By identifying the nature and extent of harms experienced by victims of different crime types, we can prioritise those crimes which are most prevalent and most harmful. A crime harm metric could also be used to measure success or failure of different strategies and policies, by considering harm prevented as well as reductions in prevalence or frequency.

Early attempts to assess the seriousness of different crimes involved surveying community members (including students, police, judges and criminal justice professionals) who ranked a selected number of offences based on their relative seriousness (see Stylianou 2003 for a review). Studies have found relative consensus among different groups (Rossi & Henry 1980; Rossi et al. 1974). Despite this, there are important limitations to using these surveys as a measure of victim harm. Crucially, perceived seriousness is a function of not only the perceived consequences of a crime but also the perceived moral wrongfulness, meaning some crimes may be considered serious regardless of whether they cause any harm (Warr 1989). Additionally, this research does not rely on victims with first-hand experience but on public perceptions, which can be swayed by media portrayals of crime (Stylianou 2003).

An alternative method of measuring harm that has grown in popularity in recent years is the use of a crime harm index (van Ruitenburg & Ruiter 2022). Crime harm indexes have now been developed for multiple countries, including Australia (House & Neyroud 2018). They were developed as an alternative to relying on offence frequencies and to better represent the concentration of crime-related harm among offenders, victims or places (Sherman, Neyroud & Neyroud 2016). A crime harm index usually involves calculating a harm weight for each offence type based on the length of the sentence for that offence—typically measured in prison days—based on either sentencing guidelines or the actual sentences imposed on offenders. The central argument behind developing a crime harm index using this method is that the sentences imposed by the courts should reflect the relative severity of that offence and the harm to the community. These indexes have been used for a variety of purposes, including measuring crime trends over time, analysing crime concentration among individuals and locations, and measuring the impact of interventions (van Ruitenburg & Ruiter 2022).

It is currently not feasible to use sentencing data to develop a harm index for cybercrime. The limits of officially recorded data on cybercrime are well established (Decker 2020; Morgan et al. 2016). Most cybercrime offences are not reported to law enforcement, and only a small number of those that are reported are investigated and prosecuted (Voce & Morgan 2023). This attrition through the criminal justice system means limited data are available. Moreover, legislation often lags behind technological developments, making it difficult to distinguish online from offline crimes, or distinguish between specific forms of cybercrime, using recorded crime data.

In this chapter we describe the development of a harm index for individual victims of cybercrime. We adopt a novel approach that draws on victim self-report data collected through the Australian Cybercrime Survey. Our approach has the advantage of reflecting victims' own experiences of the range of impacts they faced as a direct result of their victimisation. This does not rely on the views of the public, officials or judges about the theoretical harm an offence causes to the community (Ransley et al. 2018).

# Method

## Survey data

The Australian Cybercrime Survey was conducted in early 2023 and surveyed 13,887 computer users aged 18 years and over about their experiences of cyber-dependent and cyber-enabled crimes (see Voce & Morgan 2023). The survey was conducted by Roy Morgan Research using existing online panels. Proportional quota sampling, a non-probability sampling method, was used to ensure the sample was broadly reflective of the spread of people living in Australia. Quotas were based on the Australian adult population stratified by age, sex and usual place of residence, derived from Australian Bureau of Statistics (2023) population data. The resulting data were weighted by age, sex and usual place of residence to ensure the data were representative of the spread of the Australian population. Additional random iterative method weights were applied to correct for higher levels of education and more frequent internet use, which is common among online panels. Further information on the sampling and weighting procedures, as well as the concordance between the survey sample and general population characteristics, are reported in Voce and Morgan (2023).

Because we were using the data to develop a harm index for cybercrime victims, we started by limiting the sample for this study to individuals who had experienced at least one cybercrime in the 12 months prior to the survey ($n$=6,469).

## Variables

### *Victimisation*

The Australian Cybercrime Survey measured four broad types of cybercrime:

- online abuse and harassment, which refers to online communication with or about an individual which may cause them emotional distress;
- malware, which refers to software specifically developed and used to harm a computer system or network;
- identity crime and misuse, which refers to incidents where a person's personal information has been obtained or used without their permission; and
- online fraud and scams, which involve intentionally deceiving someone to obtain money or something else of value, such as personal information, using a digital device or network.

We used a bottom-up approach to measuring cybercrime victimisation. Cybercrime terminology such as 'malware', 'ransomware' and 'phishing scams' may not be fully understood by members of the public. We measured each crime type using questions about the various incidents or symptoms that may indicate a respondent has been a victim of a particular form of cybercrime. This approach was adopted for all four categories of cybercrime.

Overall, nearly half of all respondents had been a victim of at least one type of cybercrime in the 12 months prior to the survey (46.6%, *n*=6,469). Over a quarter of the sample (27.0%, *n*=3,750) had been a victim of online abuse and harassment. One in five respondents had experienced signs of a malware attack (21.8%, *n*=3,031) or identity crime and misuse (20.1%, *n*=2787). Fraud and scam victimisation was less common, with eight percent of respondents (*n*=1,082) experiencing an incident in the 12 months prior to the survey. For the purposes of this study, within each of these four main categories we aggregated related incidents or symptoms to create sub-categories, which we used to develop harm scores (see Table 1). Further information about these categories is provided in the *Appendix*.

For victims who experienced multiple types of cybercrime in the past year, it is difficult to attribute exactly which harms were related to which offence. However, harms can be compared for victims who experienced only one type of cybercrime, and the harm associated with these individual offences can be extrapolated to those who experienced combinations of different cybercrime types.

### *Harm*

To measure the wider harms associated with cybercrime victimisation, respondents who had fallen victim to any form of cybercrime in the past year were asked about the consequences they had experienced in the 12 months prior to the survey as a result of victimisation. The survey asked about 34 items in total, including practical, social, health, financial and legal impacts. This list of impacts was developed through an extensive review of the quantitative and qualitative literature on cybercrime harm and refined through consultation with academics, police and government stakeholders. It is important to acknowledge that, while a range of harms were captured in the current study, this list is not exhaustive. Future iterations of the Australian Cybercrime Survey may have additional or modified items, and these changes will be considered when comparing harms across different years.

When a victim indicated that they had experienced a particular consequence of cybercrime victimisation, such as 'my work was negatively impacted' (binary yes/no), they were then asked how much of an impact that particular consequence had on them (1=minor impact, 2=moderate impact, 3=major impact). For each of the 34 possible consequences, a score was calculated to represent whether the consequence occurred and the impact that it had: 0=consequence did not occur, 1=minor impact, 2=moderate impact, 3=major impact.

It was important that we measured victims' subjective assessment of each impact, rather than simply whether a consequence had occurred. Only counting whether a consequence had occurred would obscure the impact of more serious events (such as losing a job compared with having to buy new software). Additionally, certain consequences are likely to impact victims differently—for example, social isolation may cause significant distress for one victim and be relatively inconsequential for another.

### *Help-seeking from police or ReportCyber*

To validate our harm measure we needed to establish that it was correlated with help-seeking from police. Respondents who had been a victim of cybercrime in the 12 months prior to the survey were asked whether they had sought help, advice or support from a range of sources following the most recent incident. Victims were coded as having sought help from police when they indicated that they sought help from police and/or the Australian Cyber Security Centre through the ReportCyber platform. ReportCyber is a national online system that allows individuals, small businesses and other organisations to securely report instances of cybercrime. These reports can then be forwarded to the most applicable law enforcement agency.

This approach to validating the harm scores by measuring the relationship between harm and the likelihood of reporting to law enforcement is supported by previous research, which shows that more serious cybercrime offences are more often reported (Van de Weijer, Leukfeldt & van der Zee 2021, 2020). This is true even though many victims of high-harm crimes do not report to police due to shame, fear or a belief that police will not be understanding or proactive in helping the situation (Birdsey & Snowball 2013).

## Building the harm index

Our study proceeded in several stages. First, having measured the range and extent of harms experienced by cybercrime victims across social, health, legal, practical and financial domains, and determined the severity of the impact on the victim, we developed an overall harm score that captures the harms experienced by each victim.

Data reduction techniques such as factor analysis and principal component analysis were not used for the construction of this harm score. There were not adequate correlations between the 34 harm items (the majority under 0.4, and almost all under 0.5), resulting in factor loadings that were not meaningful in either analysis. Furthermore, the purpose of the harm score is to capture the full breadth of impacts experienced by victims, rather than reducing the number of harms included and grouping them into sets of factors.

Given the seriousness or harm associated with crime incidents, as measured by the degree of bodily injury, economic loss, emotional damage, potential for harm, and perceived wrongfulness, is the strongest correlate of victim reporting (see Xie & Baumer 2019 for a review), including for cybercrime (Van de Weijer, Leukfeldt & van der Zee 2021, 2020), we then validated our harm score by measuring the relationship between our harm score and reporting to police and ReportCyber. We did this by estimating a logistic regression to establish whether the harm scores were associated with the likelihood of reporting to law enforcement. We also used this score to illustrate how the harm from cybercrime is—much like other crime types (Dudfield et al. 2017)—highly concentrated among victims.

Finally, we developed a harm score for specific types of cybercrime, demonstrating the relative impact that each type has on victims. We could only do this for victims who experienced one type of cybercrime in the previous year (see Table 1). Victims were asked about the impact of any cybercrime incidents in the past year, meaning we were unable to ascertain the level of harm from one specific crime type when victims had experienced more than one type of cybercrime. Importantly, as shown in Table 1, the likelihood of experiencing only one type of cybercrime varied according to the type of cybercrime reported by victims. For example, 94.0 percent of ransomware victims were also a victim of at least one other type of cybercrime, while this was true for 62.8 percent of respondents who said they were a victim of compromised financial accounts.

**Table 1: Cybercrime victimisation in sample, all victims and single-offence victims, by sub-category**

| | All victims in sample (*n*=6,469) | Victims of only one crime type (*n*=3,305)[a] | |
|---|---|---|---|
| | *n* | *n* | % |
| **Online scams and fraud** | | | |
| Online shopping scams | 379 | 115 | 30.3 |
| Phishing scams | 383 | 65 | 17.0 |
| Investment and betting scams | 192 | 19 | 9.9 |
| Remote access scams | 93 | 24 | 25.8 |
| **Online abuse and harassment** | | | |
| Extortion or harassment involving images or videos | 627 | 119 | 19.0 |
| Impersonation | 1,024 | 235 | 22.9 |
| Doxxing | 223 | 36 | 16.1 |
| Stealing or sharing content without consent | 412 | 51 | 12.4 |
| Stalking and harassing | 349 | 45 | 12.9 |
| Controlling and restricting behaviour | 552 | 88 | 15.9 |
| Hate speech | 355 | 46 | 13.0 |
| Cyberbullying | 678 | 122 | 18.0 |
| Unsolicited sexual content | 1,366 | 489 | 35.8 |
| **Identity crime and misuse** | | | |
| Compromise of financial accounts | 2,457 | 914 | 37.2 |
| Other identity compromise | 498 | 32 | 6.4 |
| **Malware** | | | |
| Ransomware | 331 | 20 | 6.0 |
| Data theft and extortion | 381 | 83 | 21.8 |
| Other malware | 2,653 | 802 | 30.2 |

a: Excludes four victims who were a victim of a romance scam but no other type of cybercrime. This sample was too small for analysis. Seventy respondents had been the victim of romance scams in the last year, but only four victims experienced romance scams and no other type of cybercrime. Romance scams were not combined with other scam types because there are important differences in modus operandi and potential harm

Note: See the *Appendix* for more information on which incidents are captured within these subcategories of cybercrime

# Results

## Prevalence and impact of negative consequences

Overall, 53.1 percent of cybercrime victims were negatively impacted in some way. Forty-one percent of victims reported practical impacts, 17.9 percent reported social impacts, 15.9 percent reported health-related harms, and 15.3 percent reported financial problems. Legal issues were comparatively rare (1.8%). Table 2 presents the prevalence of each consequence experienced by cybercrime victims. The most common difficulties victims encountered were difficulty knowing which information to trust online (22.0%); being less confident using the internet for personal affairs (14.5%); having to change their personal, banking and/or contact information (13.2%); having their studies negatively impacted (12.8%); and losing trust in other people (11.6%). The consequences that were most commonly rated as having a major impact were being unable to get a loan when needed (50.0%), having to change their place of residence (48.8%), having to seek medical treatment (44.6%), and having to borrow money from family and friends (42.2%). These consequences were relatively rare, affecting fewer than three percent of victims.

**Table 2: Experiences and impact of adverse consequences of cybercrime victimisation (%)**

| | Experienced consequence (%) | Impact of consequence (%)[a] | | |
|---|---|---|---|---|
| | | Minor | Moderate | Major |
| **Practical consequences** | | | | |
| Found it harder to know which information to trust online | 22.0 | 24.8 | 46.2 | 29.0 |
| Less confident using the internet for personal affairs (eg banking, purchasing items) | 14.5 | 27.7 | 48.6 | 23.6 |
| Had to change personal, banking and/or contact information | 13.2 | 40.7 | 42.8 | 16.4 |
| Studies were negatively impacted | 12.8 | 28.8 | 39.1 | 32.1 |
| Had difficulty accessing online accounts and resources (eg bank accounts, utilities, email) | 3.9 | 30.0 | 45.9 | 24.1 |
| Had problems communicating with people (eg friends, family, employer) | 2.8 | 28.7 | 41.4 | 29.8 |
| Work was negatively impacted | 2.7 | 27.1 | 45.1 | 27.8 |
| Lost important or sentimental data | 2.0 | 22.1 | 37.4 | 40.6 |
| Had to take time off work to deal with the consequences of victimisation | 2.3 | 23.3 | 43.0 | 33.7 |
| Had problems communicating or dealing with government departments | 1.8 | 16.6 | 46.2 | 37.2 |
| Had problems communicating or dealing with businesses | 1.8 | 16.9 | 57.0 | 26.1 |
| Had to change their place of residence | 1.1 | 16.6 | 34.6 | 48.8 |

**Table 2: Experiences and impact of adverse consequences of cybercrime victimisation (%) (cont.)**

| | Experienced consequence (%) | Impact of consequence (%)[a] | | |
|---|---|---|---|---|
| | | Minor | Moderate | Major |
| **Social consequences** | | | | |
| Lost trust in other people | 11.6 | 25.4 | 46.0 | 28.6 |
| Was embarrassed or their reputation was damaged | 5.0 | 31.6 | 33.9 | 34.5 |
| Became more socially isolated | 4.5 | 28.4 | 38.9 | 32.8 |
| Relationship with their partner had been negatively impacted | 2.8 | 30.7 | 41.8 | 27.5 |
| Relationships with family and friends had been negatively impacted | 2.6 | 23.0 | 41.7 | 35.3 |
| **Health consequences** | | | | |
| Experienced difficulty sleeping | 6.7 | 21.1 | 47.3 | 31.7 |
| Overall physical health and wellbeing had deteriorated | 3.8 | 19.2 | 44.3 | 36.5 |
| Experienced mental or emotional distress | 2.6 | 23.0 | 41.7 | 35.3 |
| Had to seek psychological or counselling treatment | 2.5 | 15.4 | 46.1 | 38.5 |
| Increased their consumption of alcohol | 1.6 | 22.5 | 51.7 | 25.7 |
| Increased their consumption of drugs (legal or illegal) | 1.1 | 30.1 | 38.1 | 31.8 |
| Had to seek medical treatment | 0.1 | 18.7 | 36.8 | 44.6 |
| **Financial consequences** | | | | |
| Experienced an increase in financial stress | 5.3 | 20.6 | 47.7 | 31.7 |
| Had to buy new software | 4.1 | 29.9 | 45.2 | 24.9 |
| Had to pay for computer, phone or other hardware repairs or replacement | 3.6 | 26.5 | 44.2 | 29.3 |
| Had to buy new backup data storage or data storage devices | 2.7 | 34.2 | 45.3 | 20.5 |
| Had to borrow money from family and friends | 2.6 | 24.0 | 33.7 | 42.2 |
| Was unable to get a loan when they needed one | 1.5 | 13.8 | 36.2 | 50.0 |
| Increased the amount of time and/or money they spent gambling | 1.3 | 37.7 | 35.3 | 27.0 |
| Lost their job | 0.8 | 17.3 | 41.1 | 41.6 |
| **Legal consequences** | | | | |
| Had to commence legal action | 1.3 | 27.5 | 42.3 | 30.2 |
| Had been in trouble with the police | 0.5 | 33.6 | 33.3 | 33.1 |

a: Percentage of victims who reported that they experienced that consequence

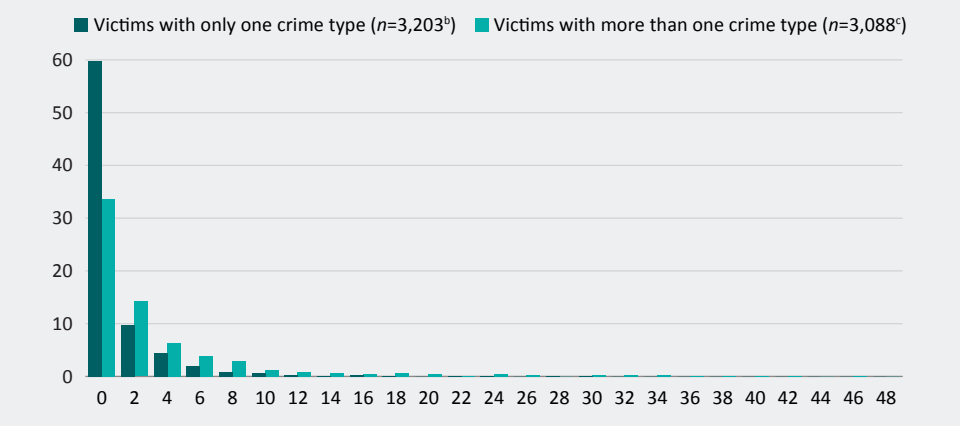
Note: *n*=6,295. Excludes 104 victims who selected did not know and 70 victims who selected prefer not to say

## Calculating harm scores

Each of the 34 items listed in Table 2 was scored according to whether the victim had experienced the consequence and how they rated its severity. Consequences were scored as: 0=consequence did not occur, 1=minor impact, 2=moderate impact, and 3=major impact. The scores for each item were then summed together to a total score, ranging from 0 to 102.

The total harm scores are displayed in Figure 1. The median harm score was 1 for all victims in the sample (range=0–99, mean=2.93). The distribution in harm scores varied according to the number of cybercrime types reported, with victimisation from more than one type of cybercrime associated with higher harm scores. Victims who only experienced one crime type in the past year had a median score of 0 (mean=1.58) while victims who experienced more than one crime type had a median harm score of 2 (mean=4.32). Irrespective of the number of crime types experienced, most victims reported no harm (47% of victims) or a harm score of between 1 and 10 (a further 47% of victims). A small number of victims experienced significant harm.

**Figure 1: Distribution of harm scores (%) (*n*=6,466[a])**




a: Excludes 4 victims who experienced romance scams only

b: Excludes 74 victims who selected 'don't know' and 29 victims who selected 'prefer not to say' to harm questions.

c: Excludes 30 victims who selected 'don't know' and 42 victims who selected 'prefer not to say' to harm questions.

Note: Figure excludes 7 victims with a harm score greater than 48

## Relationship between harm scores and help-seeking from police or ReportCyber

The next aim of the study was to validate the total harm score by analysing its relationship with seeking help from police or ReportCyber following the most recent incident. Overall, 6.9 percent of victims in the sample (*n*=227) sought help, advice or support from the police or the Australian Cyber Security Centre's online reporting portal ReportCyber following the most recent incident. We estimated a logistic regression model to determine whether the harm score was associated with the likelihood of seeking help from police, while controlling for other victim and incident characteristics that may influence help-seeking (Table 3). The sample included only victims who had experienced one cybercrime type and who had no missing data on any variables in the model (*n*=2,958).

Importantly, there was a significant relationship between a victim's total harm score and their likelihood of seeking help (adjusted odds ratio (AOR)=1.09). To illustrate this relationship, we used this model to estimate the average predictive margins, adjusted for covariates using marginal standardisation (Muller & MacLehose 2014). Predictive margins represent the average predicted probability of the outcome (help-seeking) at different values of the independent variable (the harm score). The predicted probability of reporting to police was 5.3 percent for victims with a harm score of 0, compared to 14.6 percent for victims with a harm score of 10, and 24.0 percent for victims with a harm score of 20. Based on these results we concluded that our harm score was a valid measure of the harm from cybercrime and discriminated between victims who experienced lower and higher levels of harm.

**Table 3: Victim and incident characteristics associated with seeking help from police or ReportCyber**

| | | AOR | 95% CI |
|---|---|---|---|
| Gender | Male | 1.38 | [0.99, 1.94] |
| | Female | Ref | – |
| Age | 18–34 years old | Ref | – |
| | 35–64 years | 0.47*** | [0.32, 0.70] |
| | 65 years and over | 0.86 | [0.48, 1.52] |
| First Nations | Yes | 3.04* | [1.07, 8.67] |
| | No | Ref | – |
| Small to medium enterprise status | Owner | 1.19 | [0.63, 2.26] |
| | Employee | 1.09 | [0.56, 2.12] |
| | Unaffiliated worker | 1.38 | [0.80, 2.37] |
| | Not currently working | Ref | – |
| Socio-economic dis/advantage | Q1 (most disadvantaged) | Ref | – |
| | Q2 | 1.09 | [0.64, 1.86] |
| | Q3 | 0.90 | [0.53, 1.52] |
| | Q4 (most advantaged) | 0.90 | [0.56, 1.44] |
| Partner or adult children living at home | Yes | 0.89 | [0.63, 1.27] |
| | No | Ref | – |
| Respondent disability | Yes | 1.61 | [0.96, 2.69] |
| | No | Ref | – |
| Money stolen | Yes | 1.78* | [1.16, 2.73] |
| | No | Ref | – |
| Harm score (continuous variable) | | 1.09*** | [1.05, 1.13] |
| Constant | | 0.06*** | [0.03, 0.12] |

***statistically significant at $p<0.001$, *statistically significant at $p<0.05$

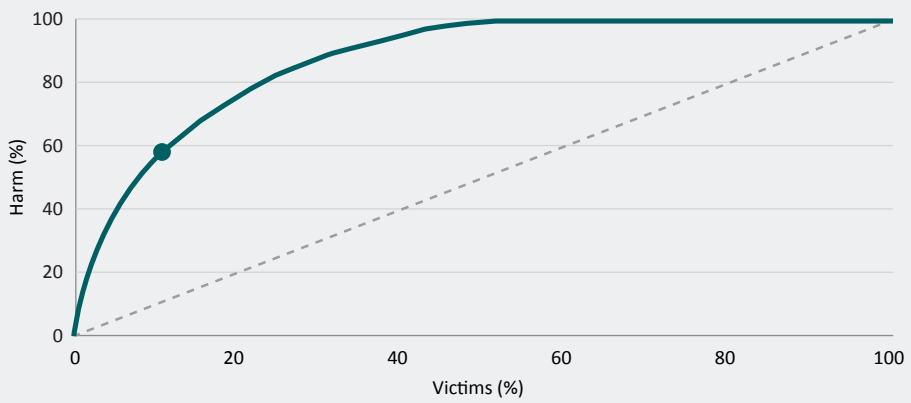
Note: AOR=adjusted odds ratio, CI=confidence interval, Ref=reference category. Hosmer–Lemeshow $\chi^2(9)=0.669$, $p=0.738$; Link test $F(2)=21.04$; $p=0.65$; Area under the receiver operating characteristic curve (AUROC)=0.66; Tolerance=0.96; Variance inflation factor (VIF)=1.04

Source: Australian Cybercrime Survey 2023

## Applying the harm score to measure concentration

Having validated the harm score, we then used it to measure the concentration of harm among cybercrime victims (Figure 2). The cumulative percentage of victims is plotted on the x-axis, and the cumulative percentage of total harm experienced by these victims is plotted on the y-axis. If harm were perfectly evenly distributed, it would follow the dotted 45-degree line. The further the actual curve is from this line, the greater the concentration of harm. Figure 2 shows that 10.9 percent of cybercrime victims accounted for around 57.7 percent of the total harm to all victims. Even if we limit this to victims who experienced at least one harm, 20.1 percent of victims accounted for the majority of reported harm. This is largely explained by repeat victims—the people who reported experiencing multiple types of cybercrime in a 12-month period.

**Figure 2: Cumulative distribution of harm among cybercrime victims**




Note: Excludes 104 victims who selected did not know and 70 victims who selected prefer not to say to harm questions

We also used this harm score to explore how the harm experienced by cybercrime victims varied according to victim characteristics. Table 4 shows the proportion of all respondents who participated in the survey, the proportion of all respondents who were victims and the corresponding proportion of the total harm score, by victim characteristics. For example, women accounted for 49.7 percent of survey respondents and 49.1 percent of cybercrime victims but 54.8 percent of all cybercrime harm.

Results show that certain groups are over-represented as victims, but that this over-representation is even more pronounced when the level of harm is taken into account. For example, respondents with a restrictive long-term health condition (ie a disability) accounted for 9.4 percent of respondents and 11.7 percent of victims but 21.6 percent of all cybercrime harm. This may be because they experience greater impacts because of their health condition, or it may be a consequence of the type of cybercrime they fall victim to. While victims who lost money in the most recent incident experienced a disproportionate level of harm (28.6% of harm but only 17.6% of victims), using the harm score also shows that most cybercrime harm (71.4%) is actually concentrated among victims who did not lose any money in the most recent incident (82.4% of victims).

**Table 4: Distribution of victims and harm, by respondent characteristics**

| | | Respondents (%) | Victims (%) | Harm (%) |
|---|---|---|---|---|
| Gender | Male | 49.9 | 50.4 | 44.2 |
| | Female | 49.7 | 49.1 | 54.8 |
| | Non-binary | 0.4 | 0.5 | 1.0 |
| Age | 18–34 years old | 29.3 | 32.4 | 38.5 |
| | 35–64 years | 48.7 | 46.2 | 45.4 |
| | 65 years and over | 22.1 | 21.4 | 16.1 |
| First Nations | Yes | 3.5 | 5.3 | 7.4 |
| | No | 96.5 | 94.7 | 92.6 |
| Lesbian, gay, bisexual or other non-heterosexual orientation | Yes | 7.9 | 9.6 | 15.1 |
| | No | 92.1 | 90.4 | 84.9 |
| Speaks a language other than English most often at home | Yes | 4.6 | 5.2 | 5.7 |
| | No | 95.4 | 94.9 | 94.3 |
| Restrictive long-term health condition | Yes | 9.4 | 11.7 | 21.6 |
| | No | 90.6 | 88.3 | 78.4 |
| Socioeconomic dis/advantage | Q1 (most disadvantaged) | 17.1 | 17.1 | 19.0 |
| | Q2 | 21.1 | 21.2 | 20.0 |
| | Q3 | 25.6 | 25.5 | 25.6 |
| | Q4 (most advantaged) | 36.2 | 36.1 | 35.4 |
| Small to medium enterprise status | Owner | 13.1 | 15.9 | 20.0 |
| | Employee | 17.2 | 17.8 | 18.5 |
| | Unaffiliated | 34.0 | 32.2 | 27.2 |
| | Not working | 35.7 | 34.1 | 34.3 |
| Lost money in most recent incident | Yes | 8.1 | 17.6 | 28.6 |
| | No[a] | 91.9 | 82.4 | 71.4 |
| Number of cybercrime types experienced | None | 53.4 | – | – |
| | One | 26.5 | 56.8 | 35.5 |
| | Two | 12.5 | 26.9 | 30.0 |
| | Three | 5.2 | 11.1 | 21.5 |
| | Four | 2.4 | 5.2 | 13.1 |

a: For respondents, this includes individuals who were not victims

## Harm index for different types of cybercrime

Having measured harm among cybercrime victims generally, and validating this score, we then developed a harm index for different types of cybercrime. This measures the relative impact of different types of cybercrime. To create the final harm index for each crime type, the prevalence of any harm within each crime type was multiplied by the mean victim total harm score for that crime type (Table 5). We then standardised the harm score so that the crime type with the lowest score—data theft and extortion—was equal to one.

**Table 5: Calculation of harm index values**

| | Prevalence of any harm (%) | Mean total harm score (range=1–102) | Prevalence of any harm × mean harm score | Standardised harm score |
|---|---|---|---|---|
| Data theft and extortion | 22.5 | 2.37 | 0.53 | 1.00 |
| Hate speech | 23.6 | 3.12 | 0.74 | 1.21 |
| Unsolicited sexual content | 24.8 | 3.63 | 0.90 | 1.37 |
| Impersonation | 39.7 | 2.97 | 1.17 | 1.64 |
| Extortion or threats involving images or videos | 40.0 | 3.05 | 1.22 | 1.69 |
| Compromise of financial accounts | 42.9 | 3.68 | 1.58 | 2.05 |
| Other malware | 38.1 | 4.22 | 1.61 | 2.08 |
| Ransomware | 41.6 | 3.88 | 1.61 | 2.08 |
| Investment and betting scams | 41.8 | 4.38 | 1.83 | 2.30 |
| Doxxing | 61.3 | 3.09 | 1.89 | 2.36 |
| Phishing scams | 51.4 | 3.94 | 2.03 | 2.50 |
| Controlling and restricting behaviour | 44.6 | 4.66 | 2.07 | 2.54 |
| Other identity compromise | 51.6 | 4.16 | 2.14 | 2.61 |
| Cyberbullying | 58.6 | 3.84 | 2.25 | 2.72 |
| Online shopping scams | 63.6 | 3.88 | 2.47 | 2.94 |
| Stealing or sharing content without consent | 53.7 | 5.96 | 3.20 | 3.67 |
| Remote access scams | 52.4 | 6.85 | 3.59 | 4.06 |
| Stalking and harassing | 57.1 | 8.86 | 5.06 | 5.53 |

Note: See the *Appendix* for more information on the incidents captured in these subcategories of cybercrime

Source: Australian Cybercrime Survey 2023

Figure 3 presents all 18 crime types from most to least harmful to individual victims of cybercrime. Two of the three highest ranking cybercrimes—stalking and harassing (5.53) and stealing or sharing content without consent (3.67)—were forms of online abuse and harassment. While the likelihood of reporting any harm was similar for victims of these and other high-ranking cybercrime types, these victims experienced more serious negative impacts across a range of domains. Three of the most harmful forms of profit-motivated cybercrime were remote access scams, online shopping scams and phishing scams. In addition to being the most harmful forms of online fraud and scams, these were also the most common (see Voce & Morgan 2023).

**Figure 3: Harm index values by offence type**

Source: Australian Cybercrime Survey 2023

## Discussion

This study measured the range and extent of harms experienced by cybercrime victims, developed and validated a harm score that accounts for the prevalence and severity of consequences for victims and used this to examine how harm varies according to victim characteristics. We then used these harm scores to develop a harm index for individual victims of cybercrime, a ranking and relative measure of 18 different categories of cybercrime.

Cybercrime victims experience a range of harms arising from their victimisation. Practical, social and health impacts were the most common harms experienced, followed by financial and legal difficulties. In fact, most crime harm is concentrated among victims who do not report any financial losses, demonstrating the importance of measuring non-financial harms. Encouragingly, most victims stated that they did not experience any of the harms listed. Harm can be reduced when incidents are successfully disrupted or steps are taken to minimise the impact on victims. For example, financial institutions may intervene when they identify or are notified of a profit-motivated cybercrime incident, either preventing transactions from being processed or recovering lost funds.

However, while more than half of victims reported not having experienced any harms, that still means a significant proportion were negatively impacted. Among those who experienced consequences, most said the impact was moderate to major. Further, when we converted these reported consequences into a harm score—which we validated by showing it was associated with police help-seeking—it was clear that cybercrime harm was heavily concentrated among a relatively small group of victims. Overall, just 10.9 percent of victims accounted for 57.7 percent of the harm to all victims who completed the survey. Much of this concentration can be explained by victims experiencing more than one type of cybercrime in a 12-month period. Repeat victims who experience multiple types of cybercrime are disproportionately impacted and should be prioritised for intervention.

Our harm index for individual victims of cybercrime shows that, among 18 common categories of cybercrime, there was considerable variation in the amount of harm reported by victims. Among the most harmful cybercrimes were stalking and harassment and having images or personal information stolen or shared without consent. These are forms of serious technology-facilitated violence that often occur as part of a pattern of abuse, often involving intimate partners (Boxall & Morgan 2021). While a substantial body of evidence shows the effects of stalking on the mental health and wellbeing of victims (Blaauw et al. 2002), research has also shown that online forms of stalking are just as harmful (Sheridan & Grant 2007). Image-based abuse has been shown to cause anxiety, depression, post-traumatic stress responses and social isolation (Bates 2017). Similarly, cyberbullying was ranked as a high-harm offence among our sample of adult victims. Most research on cyberbullying comes from adolescent and younger populations (Jenaro, Flores & Frías 2018). This finding suggests that the psychosocial, affective and academic problems associated with youth cyberbullying (Tokunaga 2010; Zych, Ortega-Ruiz & Del Rey 2015) may also be experienced by adult victims. Other types of online abuse and harassment, such as extortion or threats involving images or videos, had a lower harm score. This might be because they were more likely to be one-off incidents, because they did not always involve a credible threat to the victim's safety or because they are less harmful when they occur in isolation from other cybercrime types.

Remote access scams and online shopping scams were the two highest ranked profit-motivated cybercrimes, followed by other identity compromise. Investment scams were ranked comparatively low—despite having a higher median financial loss per report than all other scam types (Australian Competition and Consumer Commission nd)—possibly because these data include people who did not report to authorities and whose losses may not have been as significant (or, relative to their financial position, as harmful). While remote access scam reports also have a high median financial loss (Australian Competition and Consumer Commission nd), it is possible that the violation and intrusion victims feel causes more non-financial harm. Online shopping scams tend to result in much smaller financial losses, but these losses may be more acutely felt by the people who fall victim, while also impacting the perceived safety of a routine online activity.

Compromise of financial accounts was also ranked lower than other crime types. This form of identity crime and misuse includes suspicious transactions appearing on credit card statements. Voce and Morgan (2023) have shown that identity crime and misuse victims are much more likely to recover lost funds than other victims of profit-motivated cybercrime and that, when they do recover money, they recover a greater proportion of their losses. Similarly, data theft and extortion had the lowest overall harm score—probably because this includes cases of threatened and actual data theft, and many victims may not have viewed the threat as being credible.

Irrespective of the crime type, it is important to note that these harm scores are not reflective of the harms to all victims; rather, they represent the average effect and provide a way to measure harms in the aggregate. Clearly any attempt to respond to the impact on actual victims should account for their individual experience.

Importantly, individuals are only one subject of cybercrime harm, alongside the impacts on organisations, infrastructure and governments (Agrafiotis et al. 2016). This is worth considering when interpreting the harm index, which is designed for individual victims of cybercrime. For example, ransomware was ranked low on the harm index. This may be because ransomware is more often directed at small to medium enterprise owners (Voce & Morgan 2021), public services and critical infrastructure (Australian Signals Directorate 2023), where the consequences are much more serious. Given the vulnerability of small business to cybercrime, future research could use this study's methodology to develop a harm index for small business victims, accounting for harms such as disruption to everyday business, additional business expenses, impacts on reputation and revenue, loss of information, impacts on staff and legal or regulatory issues (Voce & Morgan 2023).

The most important implication of this research is the need to prevent repeat victimisation from cybercrime. While efforts to prevent cybercrime among the wider community are clearly needed, the benefits from preventing repeat victimisation are likely to be substantial and can have a meaningful impact on the overall level of cybercrime harm. While these strategies will be specific to cybercrime and, importantly, specific types of cybercrime, a wealth of evidence on the effectiveness of strategies targeting repeat victimisation can help guide these efforts (Grove et al. 2012).

This research also demonstrates the importance of timely, consistent and effective support to victims that extends beyond financial consequences. Instead, these responses should be based on an assessment of victim needs and referrals to different sources of help, whether practical, financial, health, legal or social. Failure to respond to the needs of victims exacerbates harms, undermines efforts to prevent repeat victimisation, and erodes confidence in all agencies responsible for addressing cybercrime.

Finally, we developed the harm index to help supplement other data on the prevalence and frequency of common cybercrimes. The impact on victims is rarely reflected in official reports on cybercrime (Ignatuschtschenko 2021). The cybercrime harm index can allow for monitoring of the impacts of cybercrime on victims over time, allocating victim support and law enforcement resources, and measuring the impact of prevention efforts on reduced harm to victims.

## References

*URLs correct as at October 2024*

Agrafiotis I et al. 2016. Cyber harm: Concepts, taxonomy and measurement. *Saïd Business School WP*. https://doi.org/10.2139/ssrn.2828646

Australian Bureau of Statistics (ABS) 2023. National, state and territory population, September 2022. https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/sep-2023

Australian Competition and Consumer Commission (ACCC) nd. Scam statistics. https://www.scamwatch.gov.au/research-and-resources/scam-statistics

Australian Competition and Consumer Commission 2023. *Targeting scams: Report of the ACCC on scams activity 2022*. Canberra: ACCC. https://www.accc.gov.au/about-us/publications/serial-publications/targeting-scams-reports-on-scams-activity/targeting-scams-report-of-the-accc-on-scams-activity-2022

Australian Signals Directorate (ASD) 2023. *ASD cyber threat report 2022–2023*. https://www.cyber.gov.au/about-us/view-all-content/reports-and-statistics/asd-cyber-threat-report-july-2022-june-2023

Bates S 2017. Revenge porn and mental health: A qualitative analysis of the mental health effects of revenge porn on female survivors. *Feminist Criminology* 12(1): 22–42. https://doi.org/10.1177/1557085116654565

Birdsey E & Snowball L 2013. *Reporting violence to police: A survey of victims attending domestic violence services*. Bureau Brief no. 91. Sydney: NSW Bureau of Crime Statistics and Research. https://bocsar.nsw.gov.au/media/2013/mr-bb91.html

Blaauw E, Winkel FW, Arensman E, Sheridan L & Freeve A 2002. The toll of stalking: The relationship between features of stalking and psychopathology of victims. *Journal of Interpersonal Violence* 17: 50–63. https://doi.org/10.1177/0886260502017001004

Borwell J, Jansen J & Stol W 2022. The psychological and financial impact of cybercrime victimization: A novel application of the shattered assumptions theory. *Social Science Computer Review* 40(4): 933–954. https://doi.org/10.1177/0894439320983828

Boxall H & Morgan A 2021. Intimate partner violence during the COVID-19 pandemic: A survey of women in Australia. Research report 03/2021. Sydney: Australia's National Research Organisation for Women's Safety. https://www.anrows.org.au/publication/intimate-partner-violence-during-the-covid-19-pandemic-a-survey-of-women-in-australia/

Cross C, Richards K & Smith RG 2016. The reporting experiences and support needs of victims of online fraud. *Trends & issues in crime and criminal justice* no. 518. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/ti148355

Decker E 2020. Full Count? Crime rate swings, cybercrime misses and why we don't really know the score. *Journal of National Security Law and Policy* 10: 583-604. https://jnslp.com/wp-content/uploads/2020/05/Crime-Rate-Swings-Cybercrime-Misses.pdf

Dudfield G, Angel C, Sherman LW & Torrence S 2017. The "power curve" of victim harm: Targeting the distribution of crime harm index values across all victims and repeat victims over 1 year. *Cambridge Journal of Evidence-Based Policing* 1(1): 38–58. https://doi.org/10.1007/s41887-017-0001-3

Golladay K & Holtfreter K 2017. The consequences of identity theft victimization: An examination of emotional and physical health outcomes. *Victims & Offenders* 12(5): 741–760. https://doi.org/10.1080/15564886.2016.1177766

Grove L, Farrell G, Farrington DP & Johnson SD 2012. *Preventing repeat victimisation: A systematic review*. Stockholm: Swedish National Council for Crime Prevention

House PD & Neyroud PW 2018. Developing a crime harm index for Western Australia: The WACHI. *Cambridge Journal of Evidence-Based Policing* 2(1–2): 70–94. https://doi.org/10.1007/s41887-018-0022-6

Ignatuschtschenko E 2021. Assessing harm from cyber crime. In P Cornish (ed), *The Oxford handbook of cyber security*. Oxford: Oxford University Press: chapter 8

Jansen J & Leukfeldt R 2018. Coping with cybercrime victimization: An exploratory study into the impact and change. *Journal of Qualitative Criminal Justice & Criminology* 6(2). https://doi.org/10.21428/88de04a1.976bcaf6

Jenaro C, Flores N & Frías CP 2018. Systematic review of empirical studies on cyberbullying in adults: What we know and what we should investigate. *Aggression and Violent Behavior* 38: 113–122. https://doi.org/10.1016/j.avb.2017.12.003

Morgan A, Dowling C, Brown R, Mann M, Voce I & Smith M 2016. *Evaluation of the Australian Cybercrime Online Reporting Network*. Report prepared for CrimTrac. Canberra: Australian Institute of Criminology. https://www.aic.gov.au/sites/default/files/2020-06/acorn_evaluation_report_.pdf

Muller CJ & MacLehose RF 2014. Estimating predicted probabilities from logistic regression: Different methods correspond to different target populations. *International Journal of Epidemiology* 43(3): 962–970. https://doi.org/10.1093/ije/dyu029

Ransley J et al. 2018. Developing and applying a Queensland Crime Harm Index—implications for policing serious and organised crime. In RG Smith (ed), *Organised crime research in Australia 2018*. Research Report no. 10. Canberra: Australian Institute of Criminology: 105–114. https://doi.org/10.52922/rr202682

Rossi PH & Henry JP 1980. Seriousness: A measure for all purposes? In MW Klein & KS Teilmann (eds), *Handbook of criminal justice evaluation*. Newbury Park: Sage: 489–505

Rossi PH, Waite E, Bose CE & Berk RE 1974. The seriousness of crime: Normative structure and individual differences. *American Sociological Review* 39: 224–237. https://doi.org/10.2307/2094234

Sheridan LP & Grant T 2007. Is cyberstalking different? *Psychology, Crime & Law* 13(6): 627–640. https://doi.org/10.1080/10683160701340528

Sherman L, Neyroud PW & Neyroud E 2016. The Cambridge Crime Harm Index: Measuring total harm from crime based on sentencing guidelines. *Policing: A Journal of Policy and Practice* 10(3): 171–183. https://doi.org/10.1093/police/paw003

Stylianou S 2003. Measuring crime seriousness perceptions: What have we learned and what else do we want to know. *Journal of Criminal Justice* 31(1): 37–56. https://doi.org/10.1016/s0047-2352(02)00198-8

Teunissen C, Voce I & Smith RG 2021. *Estimating the cost of pure cybercrime to Australian individuals*. Statistical Bulletin no. 34. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sb78269

Tokunaga RS 2010. Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior* 26(3): 277–287. https://doi.org/10.1016/j.chb.2009.11.014

Van de Weijer S, Leukfeldt R & Van der Zee S 2021. Cybercrime reporting behaviors among small- and medium-sized enterprises in the Netherlands. In M Weulen Kranenbarg & R Leukfeldt (eds), *Cybercrime in context: Crime and justice in digital society*. https://doi.org/10.1007/978-3-030-60527-8_17

Van de Weijer S, Leukfeldt R & van der Zee S 2020. Reporting cybercrime victimization: Determinants, motives, and previous experiences. *International Journal of Police Strategies & Management* 43(1): 17–34. https://doi.org/10.1108/PIJPSM-07-2019-0122

Van Ruitenburg T & Ruiter S 2022. The adoption of a crime harm index: A scoping literature review. *Police Practice and Research* 24(4): 423–445. https://doi.org/10.1080/15614263.2022.2125873

Voce I & Morgan A 2023. *Cybercrime in Australia 2023*. Statistical Report no. 43. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sr77031

Voce I & Morgan A 2021. *Ransomware victimisation among Australian computer users*. Statistical Bulletin no. 35. Canberra: Australian Institute of Criminology. https://doi.org/10.52922/sb78382

Warr M 1989. What is the perceived seriousness of crimes? *Criminology* 27: 795–814. https://doi.org/10.1111/j.1745-9125.1989.tb01055.x

Whitty M & Buchanan T 2015. The online dating romance scam: They psychological impact on victims—both financial and non-financial. *Criminology and Criminal Justice* 13(1): 37–61. https://doi.org/10.1177/1748895815603773

Xie M & Baumer EP 2019. Crime victims' decisions to call the police: Past research and new directions. *Annual Review of Criminology* 2(1): 217–240. https://doi.org/10.1146/annurev-criminol-011518-024748

Zych I, Ortega-Ruiz R & Del Rey R 2015. Systematic review of theoretical studies on bullying and cyberbullying: Facts, knowledge, prevention, and intervention. *Aggression and Violent Behavior* 23: 1–21. https://doi.org/10.1016/j.avb.2015.10.001

This work was originally published as *Trends & issues in crime and criminal justice* no. 706. It is available online at **https://doi.org/10.52922/ti77666**.

# Index

## A

Aboriginal and Torres Strait Islander peoples *see* Indigenous justice
acquisitive crime, 34, 39
adolescents *see* children
adult-based platforms, underage use of, 3, 178–86
    offline experiences, 179, 182–83, 185–86
    online experiences, 178, 181–82, 183–84
AI *see* artificial intelligence
alcohol use *see* drug and alcohol use
anti-lockdown and anti-vaccination protests, 2, 48, 55, 57, 58, 59, 60–61, 62
artificial intelligence, 98–109
    and child sexual abuse offending, 2, 98, 101, 105–6, 107, 108
    and child sexual abuse prevention and disruption, 2, 99, 101, 103–7, 108, 109, 130–31, 138, 139
    and child sexual abuse research, 101–2, 105–6, 107–8
AUSTRAC (Australian Transaction Reports and Analysis Centre), 128, 129, 138
Australian Criminal Intelligence Commission, 128, 129
Australian Cybercrime Survey, 204, 206, 215, 221, 222

## B

blackmail *see* sexual extortion
BYOD (bring your own device), 212, 216

## C

cannabis, 39, 44
child protection, 1, 19–29, 33
child sexual abuse material, 84, 85; *see also* sexual extortion
    and artificial intelligence, 2, 98–109, 130–31, 138, 139
    characteristics and behaviours of those who viewed, 2, 88–94
    drivers and deterrents of offending, 2, 113–24
    live streaming, 2, 127–39
    overlap with viewing fringe or radical content, 2, 84–94
    prevention, 2, 93–94, 99, 101, 103–7, 108, 109, 130–31, 138, 139
children
    of mothers in custody, 1, 19–20, 24–29
    sexual extortion impacts, 3, 142, 164–75
    sexual extortion prevalence and nature, 3, 142–61, 164
    underage use of adult-based platforms, 3, 178–86
climate change action, support for, 48, 55, 62, 63
Co-Located Caseworker Program (NSW), 1, 20–29, 33
    outcomes, 1, 27–29
    participant demographics, 23–24, 27
    participant experiences and views, 25–27
    referrals to, 23, 25, 27

cocaine, 35, 37
colonisation, 4, 5
conspiratorial beliefs, 2, 48, 49, 51–52, 59–61, 63
Convenors (restorative justice), 3, 192, 193, 196–97, 198, 200, 201
correctional settings, 1, 4, 15, 19–20, 23, 27
Corrective Services NSW, 20, 21, 26, 27, 29
COVID-19 pandemic protests, 2, 48, 55, 57, 58, 59, 60–61, 62
crime harm indexes, 221; *see also* harm index for cybercrime victims
cryptocurrency, 139
cyberbullying, 233
cybercrime; *see also* sexual extortion
    Australian Cybercrime Survey, 204, 206, 215, 221, 222
    harm index, 3, 220–34
    help-seeking, 223, 227–28
    prevention and mitigation, 215, 216, 232, 234
    ransomware, 3, 204–16, 224, 234
cybersecurity, 204, 215, 216

## D

darknet, 2, 98, 105, 113–24, 127
data breaches, 205, 210, 215
data theft and extortion, 233
dating platforms, 147, 158, 178, 179; *see also* adult-based platforms, underage use of
deaths in custody, Indigenous, 48
Department of Communities and Justice (NSW), 20, 27, 29
digitally manipulated material, 3, 149, 156, 157, 160
Discord (platform), 90, 152
disinformation, 63
disruptive protests *see* protests
domestic and family violence, 3, 191–201
double extortion, 205, 216
drug and alcohol use
    methamphetamine, 1, 38, 39, 42, 44, 45
    relationship with criminal offending, 1, 34–45
    underage users of adult-based platforms, 179, 183, 186
drug possession and selling, 39, 45
Drug Use Monitoring in Australia program, 1, 34–45

## E

ecstasy, 39, 40, 44
emotional wellbeing *see* mental health
encrypted communications, 90, 160
environmental protests, 2, 48, 55, 57, 59, 60, 62, 63
extortion, double, 205, 216
extortion, sexual *see* sexual extortion

## F

Facebook, 76, 77, 79, 90, 152, 153
false positives, 1, 2, 13, 14, 133, 138, 139
family placements, 27, 28
family violence, 3, 191–201
female prisoners, 19; *see also* mothers in custody
financial risk indicators of child sexual abuse live streaming, 2, 127–39
First Nations peoples *see* Indigenous justice
fringe or radical content, 2, 68–81, 84–94
    characteristics of those who accessed or shared, 73–74, 80–81, 88–92, 93
    exposure to, 71–73, 75, 77–81, 84, 88
    overlap with viewing child sexual abuse material, 84–94
    participation in online groups promoting, 77–79, 81, 93
    platforms used to access and share, 76–77, 79, 80, 81, 90–91
    prevention and intervention, 93–94
    sharing of, 71, 73–74, 76, 77, 79, 80, 81, 93
    violent content, 71, 80

## G

Gab (platform), 76
grooming, 164, 184

## H

harassment, 3, 232, 233
harm index for cybercrime victims, 3, 220–34
    concentration of harms, 229–30, 233
    help-seeking, 223, 227–28
    prevalence and impact of harms, 220, 222–23, 225–26, 231–33
    repeat victimisation, 229, 233, 234
    by types of cybercrime, 231–32, 233
help-seeking from police, 173, 175, 223, 227–28
heroin, 38, 39, 40, 44

## I

identity crime and misuse, 222, 233
illicit drug use *see* drug and alcohol use
iMessage, 153
incarcerated mothers *see* mothers in custody
Indigenous justice
    deaths in custody, 48
    mothers in custody, 1, 19–29, 33
    over-representation in criminal justice system, 4, 15, 19, 27
    recidivism rates, 10–14, 18
    risk assessments of Aboriginal prisoners, 1, 4–15
Instagram, 152, 153
investment scams, 233

## L

Level of Service Inventory-Revised (LSI-R), 5, 14
live streaming of child sexual abuse, 127–39
    financial risk indicators of, 2, 138–39

## M

machine learning, 99, 130–31, 138, 139; *see also* artificial intelligence
male prisoners, 1, 4–15
    recidivism rates, 10–14, 18
    Violence Risk Scale scores, 8–10, 12–14
malware, 215, 216, 222
mental health
    cybercrime harm index, 233
    mothers in custody and their children, 24, 26, 27, 28
    relationship with viewing child sexual abuse material and fringe or radical content, 2, 92, 93, 94
    sexual extortion impacts, 142, 170, 171, 174
messaging platforms, 2, 76–77, 80, 85, 154, 160
Messenger (platform), 76, 77, 81
methamphetamine, 1, 38, 39, 42, 44, 45
minors *see* children
mothers in custody, 19–20
    First Nations women, 19, 23, 24, 26, 27, 28
    NSW Co-Located Caseworker Program, 1, 20–29, 33
    pregnant mothers, 27, 28

## N

National Wastewater Drug Monitoring Program, 37, 44
New South Wales Co-Located Caseworker Program *see* Co-Located Caseworker Program (NSW)
newborn removals, 27, 28

## O

Omegle (platform), 152
ongoing victimisation, 170, 172, 174; *see also* repeat victimisation
online child sexual abuse material *see* child sexual abuse material
online fetish communities *see* adult-based platforms, underage use of
online fraud and scams, 3, 220, 222, 232, 233
online fringe or radical content *see* fringe or radical content
online sexual exploitation of children *see* child sexual abuse material; sexual extortion
online shopping scams, 232, 233
organised crime, 143, 159
out-of-home care, 19, 20, 24, 28

## P

Parler (platform), 76
peaceful protests *see* protests
Philippines, 2, 127, 128, 129, 131, 138
phishing scams, 232
police detainees, 1, 34–45
police help-seeking, 173, 175, 223, 227–28
political change, 56, 63
poly-victimisation, 170, 172, 174
pregnant women in custody, 27, 28; *see also* mothers in custody
prisoners *see* male prisoners; mothers in custody
prisons *see* correctional settings
privacy settings, social media, 147, 158, 160
profit-motivated cybercrime, 232, 233; *see also* sexual extortion
protests, 2, 48–63
    anti-lockdown and anti-vaccination, 2, 48, 55, 57, 58, 59, 60–61, 62
    environmental, 2, 48, 55, 57, 59, 60, 62, 63
    influence of conspiratorial beliefs on community support, 2, 48, 49, 51–52, 59–61, 63
    influence of protest causes on community support, 55–56, 57–58, 62
    influence of protest tactics on community support, 2, 49, 56, 57–58, 62–63
    sovereign citizen, 2, 48, 55, 57–58, 59, 60–61, 62
    violent, 2, 49, 56–58, 59–61, 62–63
psychological wellbeing *see* mental health

## R

radical content *see* fringe or radical content
radicalisation, 2, 68, 81, 84
random forest models, 130–31, 132–34, 138–39
ransomware victimisation, 3, 204–16, 224, 234
    impacts, 205, 210–13, 216
    prevention and mitigation, 215, 216
    ransom messages, 204, 208, 214, 215
    ransom payments, 205, 210, 214, 215, 216
    repeat victimisation, 205, 210, 215
    vulnerabilities, 205, 209–10, 215, 216
rapid evidence assessments, 100–109
recidivism, 1, 6, 7, 10–14, 18
Reddit, 90
remote access scams, 3, 232, 233
remunerative matchmaking platforms *see* adult-based platforms, underage use of
repeat victimisation
    cybercrime harm index, 229, 233, 234
    ongoing victimisation and poly-victimisation, 170, 172, 174
    ransomware, 205, 210, 215
    sexual extortion, 155, 159, 170, 172, 174, 175
ReportCyber, 223, 227–28
reporting of sexual offences, 114, 161, 164, 173, 175
restoration (child protection), 26, 29

restorative justice, 3, 191–201
    collaboration with external supports, 192, 199–200, 201
    Convenors, 3, 192, 193, 196–97, 198, 200, 201
    preparation and risk, 192, 193, 198–99, 201
    victim-centred processes, 192, 197–98, 201
Restorative Justice Scheme (ACT), 192–93
revictimisation *see* repeat victimisation
risk assessment, 4–5, 15, 192, 193, 198–99, 201; *see also* Violence Risk Scale
risk indicators
    child sexual abuse live streaming, 2, 127–39
    radicalisation, 68, 84
    sexual extortion, 160
risk-need-responsivity model, 4
risk-taking activities, 3
    offline, 179, 182–83, 185–86
    online, 142, 160, 178, 181–82, 183–84
Rumble (platform), 76, 81

## S

Safety by Design, 160, 175
scams, 3, 220, 222, 232, 233
sexting, 181, 183–84
sexual extortion, 3
    age of victims, 3, 146, 148, 155
    definitions, 142, 164, 169, 173
    digitally manipulated material, 3, 149, 156, 157, 160
    emotional and psychological wellbeing impacts, 142, 170, 171, 174
    financial impacts, 170, 172, 174
    prevalence and nature, 3, 142–61, 164
    prevention and intervention, 160–61, 174–75
    repeat victimisation, 155, 159, 170, 172, 174, 175
    reporting and help-seeking, 161, 164, 173, 175
    socio-environmental impacts, 142, 170, 171–72, 174
    victim-offender overlap, 170, 173, 174
sexual violence, 3, 114, 164, 191–201
Signal (platform), 90
small to medium enterprises (SMEs), 3, 204–16, 234
Snapchat, 152, 153, 181, 183
social change, 56, 63
social media platforms
    fringe or radical content, 2, 68, 72, 76–77, 79, 80, 81, 84, 90, 93
    privacy settings, 147, 158, 160
    sexual extortion, 147, 152–54, 158, 160, 175
sovereign citizen protests, 2, 48, 55, 57–58, 59, 60–61, 62
stalking, 3, 232, 233
subscription-based content-sharing platforms *see* adult-based platforms, underage use of
substance use *see* drug and alcohol use
systematic reviews, 100–109, 165–75

## T

Telegram (platform), 76, 81, 90
therapeutic support programs, 94, 124, 199–200, 201
third-party data breaches, 205, 210, 215
TikTok, 76, 81, 90, 93, 152
Torres Strait Islander peoples *see* Indigenous justice
Twitter, 90, 105

## U

undetected child sexual offenders, 2, 113, 114, 123; *see also* child sexual abuse material
unmoderated platforms, 80, 81, 84, 90

## V

victims; *see also* repeat victimisation
    cybercrime harm index, 3, 220–34
    domestic, family and sexual violence victim-survivors, 3, 191–201
    ransomware victimisation, 3, 204–16, 224, 234
    sexual extortion victimisation, 3, 142–61, 164–75
Violence Risk Scale, 1, 4, 5–15, 18
    predictive validity, 7, 12–15
    recidivism rates by risk levels, 10–11, 18
    score profile of Australian male prisoners, 8–10, 12–14
violent content *see* fringe or radical content
violent extremism, 2, 49, 68; *see also* conspiratorial beliefs
violent protests *see* protests
violent recidivism, 7, 10–14, 18

## W

WeChat, 81, 90
wellbeing *see* mental health
WhatsApp, 90, 153
women in custody, 19; *see also* mothers in custody
working from home, 216

## X

X (platform), 90, 105

**The Australian Institute of Criminology is Australia's national research and knowledge centre on crime and justice.**

© Australian Institute of Criminology 2026

**aic.gov.au**