About As Good As A Coin Toss

A Critical Criminological Analysis of Behavioural Deception Detection.



Omxus Criminology Perth WA 2026

Table of Contents

TABLE OF CONTENTS.....ERROR! BOOKMARK NOT DEFINED.

ABSTRACT..... **2**

1. INTRODUCTION..... **4**

2. THEORETICAL FRAMEWORK..... **6**

2.1 Critical Criminology and Institutional Harm 6
2.2 Labelling Theory and the Construction of "Deceptiveness" 6
2.3 Epistemic Injustice 7
2.4 Desistance Theory 8
2.5 Therapeutic Jurisprudence 9

3. METHODOLOGY..... **10**

3.1 Research Design.....10
3.2 Search Strategy 10
3.3 Source Selection Criteria 11
3.4 Limitations.....12

4. THE EMPIRICAL BASIS OF CREDIBILITY ASSESSMENT..... **14**

4.1 Accuracy of Human Deception Detection 14
4.2 Professional Overconfidence and Investigator Bias.....15
4.3 The Validity of Behavioural Cues 16
4.4 The Cue Overlap Problem: This Paper's Central Contribution 17

5. TRAUMA RESPONSE, NEURODIVERSITY, AND THE STRUCTURAL MISATTRIBUTION MECHANISM 18

5.1 The Neurobiology of Trauma Response 18
5.2 Trauma-Fragmented Memory and Narrative Credibility 19
5.3 Adverse Childhood Experiences in Criminal Justice Populations 20
5.4 Neurodiversity and Compounding Misattribution.....21

6. COMPARATIVE POLICY ANALYSIS: REFORM, REVERSAL, AND OUTCOMES **23**

6.1 Rationale and Case Selection 23
6.2 The Reform Period (2018–2021): Intervention and Outcomes 23
6.3 The Policy Reversal (2023–Present).....24
6.4 Attribution and the Limits of Causal Inference 25
6.5 The Nordic Benchmark 26

7. AUSTRALIAN EVIDENCE: EXISTING PROGRAMS WITH DOCUMENTED OUTCOMES..... **28**

8. POLICY IMPLICATIONS..... **30**

8.1 Credibility Assessment and Evidentiary Standards.....30
8.2 Trauma-Informed Interview Protocols 30
8.3 Trauma-Informed Correctional Practice.....31
8.4 The Fragility of Policy-Only Reform 31
8.5 Indigenous-Led Approaches.....31

9. CONCLUSION..... **33**

REFERENCES **35**

Abstract

This paper presents a systematic narrative review and comparative policy analysis examining the empirical validity of behavioural credibility assessment as practiced within the Australian criminal justice system. Situated within the theoretical traditions of critical criminology (Scraton, 2007; Hillyard & Tombs, 2007), labelling theory (Becker, 1963; Lemert, 1967), and epistemic injustice (Fricker, 2007), the paper interrogates the assumption that criminal justice professionals can reliably distinguish truthful from deceptive behaviour through behavioural observation. Meta-analytic evidence (Bond & DePaulo, 2006; N = 24,483 judgments across 206 studies) establishes that the accuracy of human deception detection is 54.1%, marginally exceeding chance (50%). A further meta-analysis (DePaulo et al., 2003; 158 cues across 116 studies) found that the majority of behavioural cues commonly associated with deception lack empirical support.

The paper's central scholarly contribution is the explicit integration of polyvagal theory (Porges, 2011) with the deception detection literature to articulate a mechanism by which the autonomic nervous system responses produced by acute trauma — silence, withdrawal, gaze aversion, psychomotor shutdown — are functionally indistinguishable from the cues that credibility assessors interpret as deceptive. This integration is formalised as what this paper terms a **structural misattribution mechanism**: a systematic, non-random error pattern embedded in the methodology itself, producing differential harm against the traumatised, neurodiverse, and socially marginalised populations who constitute the overwhelming majority of criminal justice subjects (Ford et al., 2020; Felitti et al., 1998). The paper presents a comparative policy analysis of New Zealand's reform and reversal period (2018–2026) as an illustrative case study, supplemented by Nordic and Australian evidence, with methodological caveats. Policy implications for the Australian context are discussed.

Keywords: credibility assessment, deception detection, trauma-informed justice, critical criminology, labelling theory, epistemic injustice, polyvagal theory, recidivism, evidence-based policy, adverse childhood experiences

1. Introduction

The Australian criminal justice system costs approximately AUD $32 billion annually (Productivity Commission, 2024) and produces a national recidivism rate of approximately 45% (Australian Institute of Health and Welfare [AIHW], 2023). This recidivism rate has remained largely stable despite decades of incremental increases in sentencing severity, policing resources, and correctional expenditure — a pattern consistent with what Garland (2001) describes as the “culture of control,” in which the escalation of punitive measures continues independently of evidence about their effectiveness.

At the centre of this system sits a methodology that has never been subjected to the evidentiary standards applied to other forensic tools: behavioural credibility assessment. Unlike DNA analysis, fingerprint matching, or toxicology — all of which must meet admissibility thresholds for scientific reliability (see *Daubert v. Merrell Dow Pharmaceuticals*, 1993, and its Australian equivalents) — behavioural credibility assessment operates without standardised protocols, without demonstrated reliability, and without disclosure of its known accuracy limitations to the triers of fact who rely upon it. Police officers assess credibility during interviews. Prosecutors rely on credibility assessments when making charging decisions. Judges evaluate demeanour in proceedings. Juries are directed to assess the credibility of witnesses as a central function of their role.

This paper investigates four research questions:

1. What does the peer-reviewed evidence establish about the accuracy of human behavioural deception detection?
2. To what extent do the behavioural presentations of acute trauma overlap with the cues that credibility assessors interpret as deceptive?
3. What are the implications of this overlap for justice outcomes, given the documented prevalence of trauma histories in criminal justice populations?
4. What does comparative policy evidence suggest about the effectiveness of reform approaches that address these structural vulnerabilities?

The paper's original contribution lies in the explicit integration and formal articulation of the relationship between polyvagal theory (Porges, 2011) and the deception detection literature, producing what is termed a **structural misattribution mechanism** — a conceptual framework for understanding how credibility assessment systematically converts trauma responses into negative credibility judgments. This mechanism is not a product of individual assessor incompetence or bias; it is an inherent limitation of a methodology that relies on behavioural cues that are empirically unrelated to deception but are empirically associated with autonomic trauma response.

The paper proceeds in nine sections. Section 2 presents the theoretical framework. Section 3 details the methodology and its limitations. Section 4 reviews the deception detection literature. Section 5 examines the intersection of trauma, neurodiversity, and credibility misattribution. Section 6 presents comparative policy evidence. Section 7 reviews Australian program evidence. Section 8 discusses policy implications. Section 9 presents the conclusion.

2. Theoretical Framework

2.1 Critical Criminology and Institutional Harm

This paper is situated within the tradition of critical criminology, which examines how criminal justice institutions produce and reproduce social harm through their ordinary operations — not only through exceptional failures but through the routine application of their standard procedures (DeKeseredy, 2011; Scraton, 2007; Taylor et al., 1973).

Critical criminology asks not merely "does the system work?" but "whom does the system work for, and at whose expense?" (Taylor et al., 1973). Hillyard and Tombs (2007) extend this analysis through their zemiological framework, arguing that criminology's exclusive focus on legally defined "crime" systematically ignores the greater harms caused by state and institutional actors — including the harms produced by the justice system's own operations.

Applied to credibility assessment, a critical lens directs attention to the structural question: if the methodology is empirically unreliable, and if its errors are not randomly distributed but systematically concentrated among traumatised, neurodiverse, and socially marginalised populations, then the methodology functions as what Scraton (2007) describes as "institutional violence" — harm produced not by individual malice but by the routine operation of institutional processes that are assumed to be neutral but are structurally biased.

This framework is applied throughout the paper: in Section 4, the deception detection evidence is read not merely as a measurement problem but as a structural mechanism through which institutional harm is produced; in Section 5, the differential impact on traumatised populations is analysed as a form of systematic, non-random error; and in Section 6, comparative policy evidence is interpreted as demonstrating that the harms produced by the punitive model are the product of policy choices that can be changed.

2.2 Labelling Theory and the Construction of "Deceptiveness"

The paper also draws on labelling theory (Becker, 1963; Lemert, 1967), which holds that deviance is not an inherent property of an act but is constructed through the

process of social reaction — specifically, through the application of labels by institutional actors with the power to define and enforce social categories.

Applied to credibility assessment, labelling theory provides a framework for understanding how the label “deceptive” or “uncooperative” is applied to individuals whose behaviour — silence, withdrawal, emotional dysregulation, inconsistent narrative — is a product of trauma rather than deceit. Once applied, this label cascades through the justice process: it influences charging decisions, bail determinations, clinical assessments, and sentencing outcomes.

Lemert (1967) distinguishes between primary deviance (the initial act) and secondary deviance (deviance produced by the institutional response to the label). In the credibility assessment context, the institutional misreading of trauma as deception constitutes a labelling event that can produce secondary deviance: the individual is processed through a system that compounds their trauma, increases their contact with criminal justice institutions, and embeds them further in the very populations that the system identifies as requiring intervention. This process is consistent with what Pratt (2007) identifies as “penal populism” — the escalation of punitive responses driven by political and media dynamics rather than evidence about what reduces harm.

2.3 Epistemic Injustice

The concept of epistemic injustice (Fricker, 2007) provides a further theoretical lens for this analysis. Fricker identifies “testimonial injustice” as a condition in which prejudice — whether conscious or structural — causes a hearer to give deflated credibility to a speaker’s testimony. Fricker’s foundational example is the case of Duwayne Brooks, witness to the murder of Stephen Lawrence, whose testimony was systematically dismissed by police — not because his account was unreliable but because structural prejudice reduced the credibility he was afforded.

Applied to credibility assessment, the epistemic injustice framework illuminates how the structural misattribution mechanism operates: individuals whose trauma responses cause them to present as “non-credible” within the assessment framework suffer a deflation of credibility that is not a product of the quality of their

testimony but of the institutional framework through which their testimony is evaluated. This is structural epistemic injustice embedded in the methodology itself.

This framework intersects with Potter's (2015) intersectional criminology, which demonstrates that interconnected identities — race, gender, class, disability — simultaneously affect experiences of criminalisation in ways that are invisible when examined through single-variable analysis. A neurodiverse woman with a trauma history who presents as silent and withdrawn during a police interview is subject to compounding credibility deflation across multiple axes simultaneously.

2.4 Desistance Theory

The paper's comparative policy analysis draws on desistance theory (Maruna, 2001; McNeill, 2006; Farrall et al., 2014), which examines the processes by which individuals cease offending. Desistance research has established that the transition away from criminal behaviour is a gradual process influenced by social bonds (Sampson & Laub, 1993), cognitive transformation (Giordano et al., 2002), identity reconstruction (Maruna, 2001), and the availability of prosocial opportunities — not by the severity of punishment.

Sampson and Laub (1993; 2016) demonstrate that social bonds developed in adulthood — employment, stable relationships, community participation — function as "turning points" that alter criminal trajectories, even for individuals with extensive prior offending. Giordano et al. (2002) extend this through their theory of cognitive transformation, identifying four shifts in the desistance process: openness to change, exposure to "hooks for change," envisioning a replacement self, and reassessing past behaviour as incompatible with the new identity. Moffitt's (1993) developmental taxonomy further establishes that the vast majority of offenders are "adolescence-limited" and desist naturally through maturation into adult social roles.

This theoretical frame is applied in Section 6: if desistance is a function of social reintegration rather than punitive incapacitation, then policies designed to facilitate reintegration should produce lower recidivism than policies designed to increase the severity of punishment. The comparative evidence from Norway, New Zealand, and Australian pilot programs is evaluated against this theoretical prediction.

2.5 Therapeutic Jurisprudence

Finally, the paper draws on therapeutic jurisprudence (Wexler & Winick, 1996), which examines how legal rules, procedures, and actors produce therapeutic or anti-therapeutic consequences for the individuals who encounter them. Therapeutic jurisprudence does not argue that therapeutic outcomes should override legal principles; rather, it holds that where legal processes can achieve their objectives while minimising anti-therapeutic consequences, they should be designed to do so.

Applied to credibility assessment, the therapeutic jurisprudence framework raises the question of whether a methodology that is 54.1% accurate and that systematically misclassifies trauma as deception is producing anti-therapeutic consequences — retraumatisation, misdiagnosis, inappropriate charging — that are not necessary for the achievement of the legal objectives it serves.

3. Methodology

3.1 Research Design

This paper employs a **systematic narrative review** of the deception detection, trauma response, and adverse childhood experiences literatures, combined with a **comparative policy analysis** of justice reform outcomes across jurisdictions.

The systematic narrative review approach (Baumeister & Leary, 1997) was selected because the research questions span multiple disciplinary domains (experimental psychology, clinical neurobiology, epidemiology, comparative public policy) and because the primary evidence consists of existing meta-analyses, systematic reviews, and evaluated program data rather than original empirical data. The paper synthesises findings across these domains to construct an integrative argument; it does not generate new primary data.

The comparative policy analysis component examines New Zealand (2018–2026) as an illustrative case study (Yin, 2018), supplemented by reference to Nordic jurisdictions and Australian pilot programs. This is not a formal comparative methodology in the sense of a structured most-similar-systems or most-different-systems design (Lijphart, 1971). New Zealand is selected as an illustrative case because it provides within-jurisdiction variation — a reform period followed by a reversal — that reduces the confounds inherent in cross-national comparison, though it does not eliminate them.

3.2 Search Strategy

Literature searches were conducted across the following databases: PsycINFO, Web of Science, Scopus, PubMed, and Google Scholar. Additional searches were conducted through the Campbell Collaboration Library (for systematic reviews of justice interventions) and government statistical repositories (NZ Department of Corrections, Australian Institute of Health and Welfare, Australian Bureau of Statistics, Productivity Commission).

Search terms were structured around four clusters corresponding to the paper's research questions: (1) deception detection terms ("deception detection," "credibility assessment," "lie detection," "behavioural cues to deception," "nonverbal deception");

(2) trauma response terms (“polyvagal theory,” “trauma response,” “tonic immobility,” “dorsal vagal,” “adverse childhood experiences,” “ACE prevalence prison”); (3) justice reform terms (“restorative justice,” “therapeutic jurisprudence,” “desistance,” “justice reinvestment,” “recidivism reduction”); and (4) comparative policy terms (“Nordic corrections,” “New Zealand corrections reform,” “penal exceptionalism”). Boolean operators combined clusters where relevant. Citation tracking was used to identify additional relevant sources.

The initial search yielded approximately 340 potentially relevant records. After screening by title and abstract and application of the inclusion and exclusion criteria detailed below, 65 sources were retained for the final synthesis. This process does not constitute a formal PRISMA-compliant systematic review; it is a structured search strategy designed to ensure comprehensive coverage of the relevant literatures within the scope of a narrative synthesis.

3.3 Source Selection Criteria

Deception detection literature (Section 4):

- Peer-reviewed meta-analyses and systematic reviews published in indexed journals
- Studies with sample sizes sufficient for meta-analytic inclusion (N > 100 judgments for individual studies; N > 1,000 for meta-analyses)
- Priority given to the most recent and comprehensive syntheses
- Supplementary inclusion of experimental studies examining professional overconfidence, cross-cultural variation, and interrogation methodology

Trauma, neurodiversity, and ACE epidemiology (Section 5):

- Peer-reviewed empirical studies and theoretical frameworks published in clinical psychology, neuroscience, forensic psychology, or public health journals
- Established theoretical models with substantial citation records
- For ACE prevalence: systematic reviews and population-level studies with prison-specific samples

- For neurodiversity: peer-reviewed studies examining the intersection of ASD, FND, and intellectual disability with criminal justice processes

Comparative policy analysis (Section 6):

- Government-published statistical data (NZ Department of Corrections, Statistics NZ, NZ Ministry of Justice)
- Evaluated program data with published outcome measures and comparison groups
- News media sources used only for policy announcements and timelines, not empirical claims

Exclusion criteria:

- Non-peer-reviewed advocacy publications excluded as primary evidence sources
- Studies with sample sizes insufficient for generalisation
- Sources relying on anecdotal evidence without systematic data collection

3.4 Limitations

This paper is subject to several methodological limitations:

5. **Secondary synthesis.** The paper relies on published meta-analyses and systematic reviews rather than conducting an independent meta-analysis or primary data collection. It therefore inherits the methodological limitations of its source studies, including their inclusion/exclusion decisions, coding procedures, and statistical approaches.
6. **Cross-disciplinary integration.** The synthesis spans experimental psychology, clinical neurobiology, epidemiology, and comparative public policy. Each field has its own methodological standards. Claims made at disciplinary boundaries should be read as integrative inferences rather than findings established within a single tradition.
7. **Absence of Australian-specific deception detection data.** The meta-analyses reviewed draw primarily on North American and European samples.

While there is no theoretical reason to expect that Australian populations would perform differently, this has not been empirically confirmed for Australian police, judges, or jurors specifically.

8. **The New Zealand comparative case.** The reform period (2018–2021) coincided with COVID-19, which independently affected both offending patterns and court processing. The pre-COVID decline establishes that the reform trend predated the pandemic, but the magnitude of the 2021 low point was influenced by COVID-related factors. The post-reversal period (2023–2026) coincides with post-pandemic economic disruption. Causal attribution is suggestive rather than definitive.
9. **Publication bias.** The evaluated Australian programs reviewed in Section 7 were selected because they have published positive outcome data. Programs attempted without positive results may be underrepresented.
10. **The structural misattribution mechanism.** The paper's central conceptual contribution is a theoretical synthesis, not an empirically tested model. While the individual components are each empirically supported, the integrative framework proposed here has not been tested as a unified model in experimental settings. This represents a limitation and an opportunity for future research.

4. The Empirical Basis of Credibility Assessment

4.1 Accuracy of Human Deception Detection

The foundational finding in deception detection research is that human beings perform poorly at distinguishing truthful from deceptive communications. This finding is consistent across a substantial body of experimental research and is not, at present, a contested claim within the relevant literature.

Bond and DePaulo (2006) conducted the most comprehensive meta-analysis of deception detection accuracy to date, synthesising 206 studies involving 24,483 individual judgments. Their principal findings:

- Mean accuracy: **54.1%** (chance = 50%; 95% CI: 53.2–55.0%)
- Truth detection accuracy: **56.1%** (a systematic truth bias)
- Lie detection accuracy: **47.4%** — below chance, indicating that assessors are less accurate at detecting lies than they would be by guessing
- No significant difference between trained professionals and untrained civilians
- Confidence was not correlated with accuracy

The below-chance performance on lie detection (47.4%) is particularly consequential for justice contexts, where the detection of deception — not the confirmation of truthfulness — is the primary operational concern.

Aamodt and Custer (2006) confirmed this in a meta-analysis of 108 studies: overall accuracy was 53.46%, with law enforcement officers performing at 55.51% — not statistically distinguishable from the general public. Judges performed at 61.64%, but this finding was based on a small number of studies and should be interpreted with caution.

Hartwig and Bond (2011) applied a lens model meta-analysis and found that individual differences in deception detection ability are “minute.” The variation in accuracy across individual assessors was far smaller than commonly assumed, suggesting that poor performance reflects a fundamental limitation of the task rather than a training deficit that can be remedied.

Hartwig et al. (2004) demonstrated this empirically: police officers achieved only chance-level accuracy (~50%) even when permitted to freely interrogate suspects, and their confidence bore no relationship to accuracy (r = 0.04).

Through the lens of critical criminology, these findings reveal something beyond a measurement problem. If credibility assessment operates at near-chance accuracy yet is treated as reliable by the institutions that deploy it, then it functions as what Hillyard and Tombs (2007) would recognise as a harm-producing mechanism embedded in institutional routine — its continued use sustained not by evidence of its effectiveness but by institutional inertia and the professional confidence of its practitioners.

4.2 Professional Overconfidence and Investigator Bias

A consistent finding is that professional training increases **confidence** without increasing **accuracy** (Vrij, 2008). **Meissner and Kassin (2002)** used signal detection theory to demonstrate that law enforcement training does not improve discrimination between truth and deception but instead produces a **lie bias** — trained officers are more likely to judge statements as deceptive, increasing false accusations without improving diagnostic accuracy. They describe this as a “double curse”: professionals are simultaneously inaccurate and confident in their inaccuracy, making them resistant to correction.

This finding has direct implications for the labelling process described in Section 2.2. If training produces a lie bias, then the application of the “deceptive” label is not a neutral diagnostic act but a systematically skewed judgment that over-identifies deception. In Becker’s (1963) terms, the label of “liar” or “uncooperative” is applied through an institutional process that is structurally biased toward its own application.

Kassin (2015) extends this analysis to interrogation methodology, demonstrating that the Reid Technique — the dominant interrogation model in Anglophone jurisdictions — is guilt-presumptive in design: it begins from the assumption that the suspect is lying and interprets all subsequent behaviour through that frame. This confirmation bias produces both false confessions and corrupted evidence chains: **Kassin et al. (2012)** found that DNA exoneration cases involving false confessions contained significantly more corroborating errors than non-confession cases,

demonstrating that the credibility misjudgment does not remain isolated but propagates through the entire evidentiary process.

Gudjonsson and Pearse (2011) note that England's adoption of the non-confrontational PEACE model — following a series of high-profile wrongful convictions attributed to coercive interrogation — demonstrates that viable alternatives to guilt-presumptive interview techniques exist and can be implemented at scale.

4.3 The Validity of Behavioural Cues

DePaulo et al. (2003) conducted the most extensive analysis of behavioural cues to deception, reviewing 158 cues across 116 studies. The findings indicate that the cues most commonly relied upon by credibility assessors lack empirical support:

- Gaze aversion: d = −0.01 (no reliable association)
- Fidgeting/restlessness: d = 0.04 (no reliable association)
- Speech hesitations: no reliable association in the hypothesised direction
- Postural shifts: d = 0.07 (no reliable association)
- Self-touching: d = 0.01 (no reliable association)

A small number of cues showed weak but statistically significant associations with deception — higher vocal pitch (d = 0.21), fewer narrative details (d = −0.30), and greater verbal tension (d = 0.26). Critically, these are not the cues that assessors rely upon.

Vrij and Granhag (2012) argue that the failure of passive behavioural observation is not a correctable training problem but a fundamental limitation of the paradigm. They propose that strategic cognitive load interviewing can produce genuine behavioural differences, but that this requires a fundamentally different approach to interviewing, not better training in the existing cue-reading framework.

The **Global Deception Research Team (2006)**, surveying 2,320 participants across 58 countries, found near-universal belief that gaze aversion indicates deception. **Taylor (2014)** further demonstrated that deception detection accuracy deteriorates in cross-cultural contexts because the behavioural indicators of deception vary across

cultural groups. In a multicultural society such as Australia, this cross-cultural instability introduces an additional source of systematic error.

4.4 The Cue Overlap Problem: This Paper's Central Contribution

The findings reviewed in Sections 4.1–4.3 establish that credibility assessment is inaccurate and that the cues it relies upon are empirically unfounded. However, the deception detection literature has not, to the author's knowledge, systematically integrated these findings with the neurobiology of trauma response to formally articulate the mechanism by which credibility assessment produces differential harm.

This paper proposes the concept of a **structural misattribution mechanism** to describe this integration. The mechanism operates as follows:

11. Credibility assessors rely on a set of behavioural cues (gaze aversion, silence, fidgeting, emotional flatness, narrative inconsistency) that they interpret as indicators of deception
12. These cues lack empirical association with deception (DePaulo et al., 2003)
13. These same cues are reliably produced by autonomic trauma response (Porges, 2011; van der Kolk, 2014) (see Section 5.1)
14. The population most frequently processed by the criminal justice system has extremely high rates of trauma exposure (Ford et al., 2020; Felitti et al., 1998) (see Section 5.3)
15. Therefore, the methodology systematically misclassifies trauma responses as deceptive behaviour, producing a non-random error pattern concentrated in the population it most frequently encounters

This is not a claim that every credibility misjudgment is trauma-related, nor that the cue overlap is the sole source of error. It is a claim that the overlap constitutes a structural vulnerability in the methodology — one that is sufficient to warrant scrutiny of the methodology's continued unregulated use.

5. Trauma Response, Neurodiversity, and the Structural Misattribution Mechanism

5.1 The Neurobiology of Trauma Response

Polyvagal theory (Porges, 2011) identifies three hierarchically organised autonomic response systems:

16. **Ventral vagal (social engagement):** The individual can communicate, modulate vocal tone, make eye contact, and engage in reciprocal social interaction. Credibility assessment presupposes that the individual being assessed is operating in this state.
17. **Sympathetic activation (fight-or-flight):** Hyperarousal — agitation, restlessness, hyper-alertness, rapid or pressured speech. In justice contexts, this may be interpreted as nervousness indicative of guilt or as aggression indicative of danger.
18. **Dorsal vagal (freeze/shutdown):** Hypoarousal — silence, immobility, dissociation, reduced eye contact, flat affect, physical withdrawal (e.g., covering face or ears). In justice contexts, this may be interpreted as evasion, disrespect, or refusal to cooperate.

The transition between these states is **autonomic** and cannot be overridden by instruction or willpower (Porges, 2011). An individual in dorsal vagal shutdown who is instructed to “look at me” or “answer the question” cannot comply — not because they are choosing not to but because the social engagement circuitry required to do so has been neurobiologically disengaged.

Neuroimaging studies provide direct visual confirmation of this mechanism. Van der Kolk’s (2014) fMRI research demonstrated that during trauma reactivation, activity in the hippocampus — the brain region responsible for organising coherent, sequential memory — is markedly suppressed, while the amygdala, the brain’s threat-detection and alarm system, shows significantly elevated activation. The result is a brain that is simultaneously flooded with alarm signals and unable to construct an ordered narrative — producing precisely the fragmented, inconsistent presentation that credibility assessors interpret as deception.

Møller et al. (2017) provide direct evidence of the prevalence and consequences of the freeze response in justice-relevant contexts: in a study of 298 women presenting at a Stockholm rape crisis centre, 70% reported significant tonic immobility during the assault and 48% reported extreme tonic immobility. This involuntary freeze response was a significant predictor of subsequent PTSD. **De la Torre Laso (2024)** documents that this same response is frequently misinterpreted within justice systems as consent or lack of resistance — a specific instance of the structural misattribution mechanism in operation.

Through the theoretical lens of epistemic injustice (Fricker, 2007), the dorsal vagal shutdown response constitutes a paradigmatic case of testimonial injustice: the individual's capacity to provide testimony is neurobiologically impaired, and this impairment is then interpreted by the institutional framework as evidence against the individual's credibility. The injustice is structural, not interpersonal — it is produced by the interaction between the individual's autonomic response and the institutional methodology through which that response is interpreted.

5.2 Trauma-Fragmented Memory and Narrative Credibility

Brewin (2011) and **Brewin et al. (1996)** establish that traumatic memories are encoded through processes that differ from ordinary episodic memory. Under extreme stress, hippocampal functioning is suppressed while amygdala activation is elevated, producing memories that are fragmented, sensory-dominant, temporally disordered, and variable across retellings.

Howe and Knott (2015) directly address the implications for judicial processes, demonstrating that the characteristics of trauma-affected memories — fragmentary recall, amnesic gaps, out-of-order sequencing, inconsistent details — are normal consequences of traumatic encoding but are routinely misinterpreted by legal decision-makers as indicators of fabrication. This is consistent with **Smith and Skinner's (2012)** courtroom observations, which found that trauma responses in sexual assault victims — emotional numbness, fragmented recall, delayed reporting — were systematically misread by legal professionals as indicators of unreliability, with gendered expectations of "ideal victim" behaviour further undermining credibility.

In labelling theory terms (Becker, 1963), the inconsistency produced by trauma-fragmented memory functions as a trigger for the application of the “unreliable” or “deceptive” label — a label that, once applied, cascades through the institutional process and produces secondary consequences independent of the individual’s actual veracity.

5.3 Adverse Childhood Experiences in Criminal Justice Populations

The structural misattribution mechanism acquires its significance from the prevalence of trauma in the population to which credibility assessment is applied.

Ford et al. (2020) conducted a systematic review and meta-analysis of ACE prevalence in prison populations, published in *Health & Justice*. Their analysis found that **84.2%** of prisoners in a Welsh sample reported at least one ACE, compared to 46.6% in the general population.

Additional peer-reviewed estimates confirm elevated ACE prevalence:

- **Messina and Grella (2006):** 57% of incarcerated women in a California sample reported four or more ACEs
- **Baglivio et al. (2014):** 97% of juvenile offenders had at least one ACE, with a mean ACE score of 4.89 (compared to a population average of approximately 1.6) — establishing a dose-response relationship between ACE accumulation and earlier onset of offending
- **Indig et al. (2011):** In an Australian study, 89% of young people in NSW custody reported exposure to traumatic events, with rates of mental health disorders and substance use dramatically exceeding general population rates
- **Felitti et al. (1998):** The foundational CDC-Kaiser ACE Study established a dose-response relationship between ACE score and multiple negative outcomes, including substance use disorder and criminal justice contact

The substance use pathway is particularly well-documented. **Dube et al. (2003)** demonstrated a graded dose-response relationship: individuals with five or more ACEs were 7–10 times more likely to report drug addiction. **Khoury et al. (2010)** confirmed that 75% of individuals presenting for substance use treatment report trauma histories.

Menzies (2019) extends this analysis to the Australian context, documenting how colonisation, forced child removals (the Stolen Generations), and ongoing structural disadvantage create intergenerational trauma pathways directly linked to Aboriginal and Torres Strait Islander overrepresentation in the criminal justice system. This is consistent with Potter's (2015) intersectional framework: Indigenous Australians experience compounding credibility deflation across axes of race, historical trauma, and socioeconomic marginalisation simultaneously.

5.4 Neurodiversity and Compounding Misattribution

Individuals with Autism Spectrum Disorder (ASD), Functional Neurological Disorder (FND), intellectual disability, and ADHD display communication patterns that compound the structural misattribution mechanism.

Crane et al. (2016) surveyed autistic adults and police officers in England and Wales, finding that only 39% of autistic individuals disclosed their diagnosis to police, and that autistic communication patterns (reduced eye contact, flat affect, literal interpretation of questions) were routinely misinterpreted by officers as evasiveness. **Cooper et al. (2022)** conducted a systematic review finding that autistic people display verbal cues associated with deception in neurotypical populations even when truthful, including statement-evidence inconsistencies and sparse detail.

Perske (2011) documented 75 cases of individuals with intellectual disabilities who falsely confessed to serious felonies, demonstrating that heightened suggestibility, desire to please authority figures, and impaired understanding of legal rights make this population acutely vulnerable to coerced confessions — a vulnerability amplified by the lie bias produced by professional interrogation training (Meissner & Kassin, 2002).

Datta and Blum (2024) review FND in forensic contexts, finding that distinguishing functional neurological symptoms from malingering poses a persistent forensic challenge and that FND symptoms can be exacerbated by the stress of legal proceedings themselves.

The intersection of trauma and neurodiversity produces what might be termed **compounding misattribution**: an individual with ASD who is simultaneously experiencing trauma reactivation presents a behavioural profile that the credibility

assessment framework is structurally unable to interpret accurately across multiple axes simultaneously. In Fricker's (2007) terms, this constitutes a compounding of testimonial injustice in which each axis of difference produces an independent credibility deflation, and the deflations interact multiplicatively rather than additively.

6. Comparative Policy Analysis: Reform, Reversal, and Outcomes

6.1 Rationale and Case Selection

New Zealand is selected as the primary comparative case for three reasons: (1) shared common-law legal system and institutional architecture with Australia; (2) comparable indigenous overrepresentation (Māori: 17.8% of population, 52–56% of prisoners; Aboriginal and Torres Strait Islander: 3.8% of population, 36% of prisoners); and (3) within-jurisdiction policy variation that reduces the confounds of cross-national comparison.

Nordic jurisdictions (Norway, Finland) are referenced as supplementary comparators representing sustained, long-term reform commitment. The comparison is informed by Pratt's (2008) analysis of "Scandinavian exceptionalism," while acknowledging Ugelvik and Dullum's (2012) critical caveat that Nordic exceptionalism should not be uncritically idealised.

6.2 The Reform Period (2018–2021): Intervention and Outcomes

Key interventions included a government commitment to 30% prison population reduction, launch of the Hōkai Rangi strategy (co-designed with Māori), expansion of Te Pae Oranga iwi community panels, Sentencing Act amendments mandating restorative justice referrals, the Misuse of Drugs Amendment Act 2019 (police discretion to divert drug-affected individuals to health services), and investment in the Alcohol and Other Drug Treatment Court (AODTC).

Outcomes (using 12-month re-imprisonment as the consistent metric):

Indicator	Pre-Reform (2018)	Reform Peak (2021)	Change
Prison population	10,820	7,702	−29%
12-month re-imprisonment	32%	24%	−8 percentage points

Indicator	Pre-Reform (2018)	Reform Peak (2021)	Change
Māori imprisonment rate (per 100,000)	657	459	−30%
Restorative justice referrals (annual)	~4,000	~12,000	+200%

COVID-19 caveat: The 2021 figure was partially influenced by pandemic-related effects. However, the population was already declining pre-COVID — from 10,820 (June 2018) to approximately 8,500 (late 2019) — a 21% reduction attributable to policy. COVID accelerated an existing trend; it did not create it.

These outcomes are consistent with the predictions of desistance theory. The evaluated programs confirm this:

- **AODTC:** All participants showed 45% less reoffending than a matched comparison group at 12 months; graduates specifically showed 86% less reoffending at 12 months (NZ Ministry of Justice, 2019a)
- **Te Pae Oranga:** 22% relative reduction in reoffending harm, with an average reduction of 26.9 equivalised prison days; statistically significant improvements across all four reconviction outcome measures at 12 months (Walton et al., 2020)

Strang et al. (2013) conducted a Campbell Collaboration systematic review of 10 RCTs (N = 1,880 offenders) finding that restorative justice conferences produced a modest but cost-effective reduction in reoffending. **Latimer et al. (2005)** confirmed in a separate meta-analysis that restorative justice programs were significantly more effective than traditional approaches across multiple outcomes.

6.3 The Policy Reversal (2023–Present)

The incoming coalition government reinstated Three Strikes sentencing, toughened penalties through the Sentencing Reform Act, paused AODTC expansion, and ended the Hōkai Rangi strategy without renewal.

Post-reversal outcomes (2025–2026):

Indicator	Reform Peak (2021)	Post-Reversal (Jan 2026)	Change
Prison population	7,702	11,000+	+43% (all-time record)
Imprisonment rate (per 100,000)	~150	187	+25%
Māori % of prison population	50%	52–56%	Increased
Remand (unconvicted) % of total	~20%	40–53%	Approximately doubled
Prisoner-on-prisoner assaults (annual)	Lower baseline	1,558	All-time record
Prisoner-on-staff assaults (annual)	Lower baseline	1,080	All-time record
Per-prisoner annual cost (NZD)	Lower	$201,408	Record high
New prison construction (announced)	None required	$3.5B+ NZD	Substantial commitment

12-month re-imprisonment: The most recent published figure is 26% — a 2 percentage point increase from the reform-period low of 24%, with the trajectory worsening.

6.4 Attribution and the Limits of Causal Inference

The NZ case provides suggestive but not definitive evidence for causal attribution. The within-jurisdiction design controls for legal system, culture, and demographics. However, COVID-19, post-pandemic economic disruption, policy implementation lag, and concurrent unrelated policy changes limit causal inference.

A relevant finding: the NZ Ministry of Justice's own briefing to ministers (February 2025) noted that the decline in violent victimisation observed in 2024–2025 "may indicate a return to a trend seen from 2018–2022" — suggesting that the crime decline predates the implementation of the punitive policies and may reflect a continuation of the reform-era trend.

The NZ evidence is most persuasive as part of a **convergent evidence pattern**: Norway (sustained reform, sustained improvement, 30+ years), Finland (sustained reform, sustained improvement, 60+ years), NZ reform period (partial reform, partial improvement), NZ reversal period (reversal, deterioration on all measured indicators), and Australian pilot programs (localised reform, localised improvement). The consistency of direction across multiple jurisdictions, timeframes, and institutional contexts strengthens the inference, though no single case study is sufficient to establish a universal causal law.

Through the lens of desistance theory, the NZ reversal is interpretable as a natural test of the theoretical prediction: when policies shift from facilitating desistance to inhibiting it, recidivism should increase and system costs should rise. The data are consistent with this prediction.

Cost-benefit context: Aos et al. (2006), in a landmark cost-benefit analysis of 545 evaluations, found that evidence-based community corrections programs cost thousands of dollars less than prison while producing equal or greater public safety benefits, with some programs returning up to $27 in benefits per dollar invested. The NZ government's current trajectory represents the inverse of this evidence-based approach.

6.5 The Nordic Benchmark

Jurisdiction	Reform Duration	Recidivism	Metric	Incarceration Rate
Norway	30+ years (sustained)	~20%	2-year reconviction	75/100,000
Finland	60+ years (sustained)	24–31%	2-year reconviction	57/100,000

Jurisdiction	Reform Duration	Recidivism	Metric	Incarceration Rate
NZ reform period	3–4 years (partial)	24%	12-month re-imprisonment	~150/100,000
NZ post-reversal	2–3 years (reversed)	26% and worsening	12-month re-imprisonment	187/100,000
Australia	No systematic reform	~45%	2-year reconviction	167/100,000

Note. The recidivism figures in this table are not directly comparable across jurisdictions because they use different metrics and time horizons. Norway, Finland, and Australia report 2-year reconviction rates, while New Zealand reports 12-month re-imprisonment rates. Reconviction at 2 years is consistently higher than re-imprisonment at 12 months across virtually all jurisdictions, meaning the NZ reform figure (24%) would likely be higher if measured on the same basis as the Nordic figures. Direct numerical comparison should therefore be avoided; the table is presented to illustrate the directional relationship between reform commitment and recidivism trajectory, not to establish precise cross-jurisdictional equivalence.

The Nordic data suggests that the duration and consistency of reform commitment moderates the magnitude and durability of outcomes. Pratt's (2008) analysis suggests that the institutional and cultural preconditions for sustained reform — welfare state infrastructure, media responsibility, depoliticisation of criminal justice — are more difficult to achieve in Anglophone countries with adversarial political systems. This represents a structural challenge for Australian reform efforts that should be acknowledged rather than minimised.

7. Australian Evidence: Existing Programs with Documented Outcomes

Program	Jurisdiction	Intervention	Key Outcome	Source
Bourke Justice Reinvestment	NSW	Community-led, Aboriginal-designed prevention and diversion	Reoffending −18%; assaults −39%; family violence −39%; $3.1M savings (2017)	Just Reinvest NSW; Schwartz et al. (2017)
Koori Courts	VIC	Aboriginal sentencing courts with Elder involvement	Lower reoffending; increased communication between offenders, magistrates, and communities	Marchetti & Daly (2007); County Court of Victoria
MERIT Program	NSW	Magistrates Early Referral Into Treatment	Reduced reoffending for drug-affected defendants	NSW Health
Drug Courts	National	Therapeutic jurisprudence, supervised treatment	Consistently lower reoffending than standard courts	BOCSAR (NSW); state evaluations
PACER	VIC	Police + mental health professional co-response	Reduced unnecessary psychiatric detention	Victoria Police

Marchetti and Daly (2007) provide the most rigorous academic evaluation of Australian Indigenous sentencing courts, finding that these courts produce increased understanding and communication between offenders, magistrates, and Indigenous

communities, with culturally appropriate processes that avoid the epistemic injustice (Fricker, 2007) inherent in adversarial proceedings conducted in a framework alien to the individuals being processed.

Schwartz et al. (2017) document that the Bourke Justice Reinvestment project — a community-led, place-based alternative to incarceration — achieved both improved justice outcomes and substantial cost savings, supporting the cost-benefit proposition established by Aos et al. (2006).

These programs share characteristics consistent with the therapeutic jurisprudence framework (Wexler & Winick, 1996): they seek to achieve legal objectives while minimising anti-therapeutic consequences; they address underlying contributors to offending (trauma, substance dependence, social disconnection) rather than relying solely on incapacitation; and they produce measurable recidivism reductions at lower per-participant costs. The limitation is scale: they exist in specific jurisdictions, are variably funded, and are not available to the majority of individuals who pass through the system.

8. Policy Implications

The evidence reviewed in this paper has several implications for Australian criminal justice policy. These are presented as implications arising from the evidence, not as advocacy positions. The distinction matters: the evidence establishes what is currently known; the appropriate policy response involves value judgments, resource allocation decisions, and democratic deliberation that extend beyond the scope of empirical evidence.

8.1 Credibility Assessment and Evidentiary Standards

The evidence reviewed in Section 4 raises questions about the evidentiary status of behavioural credibility assessment within Australian courts. If the methodology performs at 54.1% accuracy and relies on cues that lack empirical validity, its continued use without disclosure to triers of fact appears inconsistent with the evidentiary standards applied to other forensic methodologies. The principle of informed decision-making in jury trials suggests that jurors directed to assess witness credibility should be informed of the known limitations of the methodology they are applying.

8.2 Trauma-Informed Interview Protocols

The structural misattribution mechanism described in Sections 4.4 and 5.1 suggests that current police interview practices may systematically misclassify trauma responses as indicators of deception. Trauma-informed interview training — which equips officers to recognise autonomic responses and to avoid interpreting them through a deception framework — addresses this structural vulnerability. The PEACE model's adoption in England (Gudjonsson & Pearse, 2011), the QPS ISACURE programme in Queensland, and the UK College of Policing's trauma-informed framework demonstrate that implementation is feasible.

The evidence further suggests that the concept of "police complex spiral trauma" (Papazoglou & Tuttle, 2018) — cumulative traumatic exposure affecting officers' own decision-making and empathy — may compound the structural misattribution mechanism. Trauma-informed training that addresses officers' own trauma exposure

may therefore improve both officer wellbeing and the accuracy of credibility assessments.

8.3 Trauma-Informed Correctional Practice

Miller and Najavits (2012) document that trauma-informed correctional care requires integration of clinical and security goals, systematic screening protocols, and administrative support. **Levenson and Willis (2019)** demonstrate that trauma-informed care increases offender responsivity to rehabilitation programming — a finding consistent with desistance theory's emphasis on the social and relational conditions necessary for identity transformation (Giordano et al., 2002; Maruna, 2001).

8.4 The Fragility of Policy-Only Reform

The New Zealand case illustrates that reform gains achieved through departmental policy rather than legislation are vulnerable to reversal following changes of government. If reform is pursued in Australia, the NZ experience suggests that embedding reforms in primary legislation provides greater durability than administrative policy changes. This is consistent with Pratt's (2007) analysis of penal populism: in political systems where criminal justice policy is electorally weaponised, reforms that rely on political goodwill rather than legislative mandate are structurally fragile.

8.5 Indigenous-Led Approaches

Both the New Zealand (Te Pae Oranga) and Australian (Bourke, Koori Courts) evidence indicates that Indigenous-led justice approaches produce measurable improvements. The common elements are community ownership, cultural appropriateness, and relational rather than adversarial accountability. Through the lens of epistemic justice (Fricker, 2007), these approaches partially remediate the testimonial injustice inherent in processing Indigenous individuals through institutional frameworks designed without reference to their cultural communicative norms. The evidence supports continued investment in and scaling of these approaches, consistent with the principles of self-determination and community empowerment documented by Menzies (2019) and Marchetti and Daly (2007).

9. Conclusion

This paper has reviewed evidence across four domains — deception detection accuracy, trauma neurobiology, ACE epidemiology, and comparative justice policy — to examine the empirical foundations of credibility assessment as practiced in the Australian criminal justice system.

The paper's central contribution is the articulation of a **structural misattribution mechanism**: the integration of polyvagal theory (Porges, 2011) with the deception detection literature (Bond & DePaulo, 2006; DePaulo et al., 2003) to demonstrate that the behavioural cues used in credibility assessment overlap systematically with the autonomic responses produced by trauma, in a population characterised by extremely high trauma prevalence (Ford et al., 2020). Neuroimaging evidence confirms that during trauma reactivation, hippocampal activity is suppressed while amygdala activation is elevated — producing precisely the fragmented, inconsistent presentation that credibility assessors interpret as deception (van der Kolk, 2014). This mechanism produces differential harm against traumatised, neurodiverse, and socially marginalised individuals — not through individual assessor prejudice but through a structural limitation of the methodology itself.

Through the theoretical lenses of critical criminology (Scraton, 2007; Hillyard & Tombs, 2007), labelling theory (Becker, 1963; Lemert, 1967), and epistemic injustice (Fricker, 2007), this mechanism is interpretable as a form of institutional harm: a routine process assumed to be neutral that produces systematic, non-random errors with consequential justice implications.

The comparative policy evidence — from Norway, Finland, New Zealand, and Australian pilot programs — is consistent with the predictions of desistance theory (Maruna, 2001; Sampson & Laub, 1993): policies that facilitate social reintegration produce lower recidivism than policies focused on punitive incapacitation. The New Zealand case provides illustrative evidence that reform produces improvement and reversal produces deterioration within a single jurisdiction, though the methodological caveats detailed in Section 3.4 limit the strength of causal attribution.

This paper does not claim to resolve the policy questions its findings raise. The appropriate institutional response to empirically questionable methodology involves

value judgments, resource considerations, and democratic deliberation that extend beyond empirical evidence alone. What the evidence does establish is that the current approach rests on assumptions that are not supported by the available research; that alternatives with documented outcomes exist and are operating within Australia and comparable jurisdictions; and that the consequences of the current approach — measured in recidivism, institutional cost, and human impact — are substantial and measurable.

Whether that evidence is sufficient to warrant institutional change is a question this paper presents to its readers rather than answers on their behalf.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner*, 15(1), 6–11.

Australian Institute of Health and Welfare. (2023). *The health of Australia's prisoners 2022*. AIHW.

Aos, S., Miller, M., & Drake, E. (2006). *Evidence-based public policy options to reduce future prison construction, criminal justice costs, and crime rates*. Washington State Institute for Public Policy.

Baglivio, M. T., Epps, N., Swartz, K., Huq, M. S., Sheer, A., & Hardt, N. S. (2014). The prevalence of adverse childhood experiences (ACE) in the lives of juvenile offenders. *Journal of Juvenile Justice*, 3(2), 1–23.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology*, 1(3), 311–320. https://doi.org/10.1037/1089-2680.1.3.311

Becker, H. S. (1963). *Outsiders: Studies in the sociology of deviance*. Free Press.

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2

Brewin, C. R. (2011). The nature and significance of memory disturbance in posttraumatic stress disorder. *Annual Review of Clinical Psychology*, 7, 203–227. https://doi.org/10.1146/annurev-clinpsy-032210-104544

Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of posttraumatic stress disorder. *Psychological Review*, 103(4), 670–686.

Cooper, D. S., Uppal, D., Railey, K. S., Wilson, A. B., Maras, K., Zimmerman, E., Bornman, J., & Shea, L. L. (2022). Policy gaps and opportunities: A systematic review of autism spectrum disorder and criminal justice intersections. *Autism*, 26(5), 1060–1075. https://doi.org/10.1177/13623613211070341

Crane, L., Maras, K. L., Hawken, T., Mulcahy, S., & Memon, A. (2016). Experiences of autism spectrum disorder and policing in England and Wales. *Journal of Autism and Developmental Disorders*, 46(6), 2028–2041. https://doi.org/10.1007/s10803-016-2729-1

Datta, V., & Blum, A. W. (2024). Forensic assessment of somatoform and functional neurological disorders. *Behavioral Sciences & the Law*, 42(3), 163–175. https://doi.org/10.1002/bsl.2651

de la Torre Laso, J. (2024). The reality of tonic immobility in victims of sexual violence. *Journal of Interpersonal Violence*, 39(5–6), 1200–1222. https://doi.org/10.1177/15248380231191232

DeKeseredy, W. S. (2011). *Contemporary critical criminology*. Routledge.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74

Dube, S. R., Felitti, V. J., Dong, M., Chapman, D. P., Giles, W. H., & Anda, R. F. (2003). Childhood abuse, neglect, and household dysfunction and the risk of illicit drug use. *Pediatrics*, 111(3), 564–572. https://doi.org/10.1542/peds.111.3.564

Farrall, S., Hunter, B., Sharpe, G., & Calverley, A. (2014). *Criminal careers in transition: The social context of desistance from crime*. Oxford University Press.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. *American Journal of Preventive Medicine*, 14(4), 245–258. https://doi.org/10.1016/S0749-3797(98)00017-8

Ford, K., Barton, E. R., Newbury, A., Hughes, K., Bezeczky, Z., Roderick, J., & Bellis, M. A. (2020). Understanding the prevalence of adverse childhood experiences (ACEs) amongst people in prison. *Health & Justice*, 8, Article 22. https://doi.org/10.1186/s40352-020-00115-5

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

Garland, D. (2001). *The culture of control: Crime and social order in contemporary society*. University of Chicago Press.

Giordano, P. C., Cernkovich, S. A., & Rudolph, J. L. (2002). Gender, crime, and desistance: Toward a theory of cognitive transformation. *American Journal of Sociology*, 107(4), 990–1064. https://doi.org/10.1086/343191

Global Deception Research Team. (2006). A world of lies. *Journal of Cross-Cultural Psychology*, 37(1), 60–74. https://doi.org/10.1177/0022022105282295

Gudjonsson, G. H., & Pearse, J. (2011). Suspect interviews and false confessions. *Current Directions in Psychological Science*, 20(1), 33–37. https://doi.org/10.1177/0963721410396824

Hartwig, M., & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin*, 137(4), 643–659. https://doi.org/10.1037/a0023589

Hartwig, M., Granhag, P. A., Stromwall, L. A., & Vrij, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly*, 7(4), 429–456. https://doi.org/10.1177/1098611104264748

Hillyard, P., & Tombs, S. (2007). From 'crime' to social harm? *Crime, Law and Social Change*, 48(1–2), 9–25. https://doi.org/10.1007/s10611-007-9079-z

Howe, M. L., & Knott, L. M. (2015). The fallibility of memory in judicial processes. *Memory*, 23(5), 633–656. https://doi.org/10.1080/09658211.2015.1010709

Indig, D., Vecchiato, C., Haysom, L., Beilby, R., Carter, J., Champion, U., Gaskin, C., Godhino, H., Whitton, G., & Shrapnel, R. (2011). *2009 NSW Young People in Custody Health Survey: Full report*. Justice Health and Juvenile Justice.

Kassin, S. M. (2015). The social psychology of false confessions. *Social Issues and Policy Review*, 9(1), 25–51. https://doi.org/10.1111/sipr.12009

Kassin, S. M., Bogart, D., & Kerner, J. (2012). Confessions that corrupt: Evidence from the DNA exoneration case files. *Psychological Science*, 23(1), 41–45. https://doi.org/10.1177/0956797611422918

Khoury, L., Tang, Y. L., Bradley, B., Cubells, J. F., & Ressler, K. J. (2010). Substance use, childhood traumatic experience, and posttraumatic stress disorder in an urban civilian population. *Depression and Anxiety*, 27(12), 1077–1086. https://doi.org/10.1002/da.20751

Latimer, J., Dowden, C., & Muise, D. (2005). The effectiveness of restorative justice practices: A meta-analysis. *The Prison Journal*, 85(2), 127–144. https://doi.org/10.1177/0032885505276969

Lemert, E. M. (1967). *Human deviance, social problems, and social control*. Prentice-Hall.

Levenson, J. S., & Willis, G. M. (2019). Implementing trauma-informed care in correctional treatment and supervision. *Journal of Aggression, Maltreatment & Trauma*, 28(4), 481–501. https://doi.org/10.1080/10926771.2018.1531959

Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Science Review*, 65(3), 682–693.

Marchetti, E., & Daly, K. (2007). Indigenous sentencing courts: Towards a theoretical and jurisprudential model. *Sydney Law Review*, 29(3), 415–443.

Maruna, S. (2001). *Making good: How ex-convicts reform and rebuild their lives*. American Psychological Association.

McNeill, F. (2006). A desistance paradigm for offender management. *Criminology & Criminal Justice*, 6(1), 39–62. https://doi.org/10.1177/1748895806060666

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26(5), 469–480. https://doi.org/10.1023/A:1020278620751

Menzies, K. (2019). Understanding the Australian Aboriginal experience of collective, historical and intergenerational trauma. *International Social Work*, 62(6), 1522–1534. https://doi.org/10.1177/0020872819870585

Messina, N., & Grella, C. (2006). Childhood trauma and women's health outcomes in a California prison population. *American Journal of Public Health*, 96(10), 1842–1848. https://doi.org/10.2105/AJPH.2005.082016

Miller, N. A., & Najavits, L. M. (2012). Creating trauma-informed correctional care. *European Journal of Psychotraumatology*, 3(1), 17246. https://doi.org/10.3402/ejpt.v3i0.17246

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674

Møller, A., Søndergaard, H. P., & Helström, L. (2017). Tonic immobility during sexual assault — a common reaction predicting post-traumatic stress disorder and severe depression. *Acta Obstetricia et Gynecologica Scandinavica*, 96(8), 932–938. https://doi.org/10.1111/aogs.13174

New Zealand Department of Corrections. (2025). *Prison statistics — March 2025*. https://www.corrections.govt.nz/resources/statistics

New Zealand Ministry of Justice. (2019a). *AODTC Quantitative Outcomes Evaluation, June 2019*. Wellington: Ministry of Justice. https://www.justice.govt.nz/assets/AODTC-Quantitative-Outcomes-Evaluation-June-2019.pdf

New Zealand Ministry of Justice. (2025). *Justice sector projections*. https://www.justice.govt.nz

Papazoglou, K., & Tuttle, B. M. (2018). Fighting police trauma: Practical approaches to addressing psychological needs of officers. *SAGE Open*, 8(3). https://doi.org/10.1177/2158244018794794

Perske, R. (2011). Perske's list: False confessions from 75 persons with intellectual disability. *Intellectual and Developmental Disabilities*, 49(5), 365–373. https://doi.org/10.1352/1934-9556-49.5.365

Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation*. W. W. Norton.

Potter, H. (2015). *Intersectionality and criminology: Disrupting and revolutionizing studies of crime*. Routledge.

Pratt, J. (2007). *Penal populism*. Routledge.

Pratt, J. (2008). Scandinavian exceptionalism in an era of penal excess. *British Journal of Criminology*, 48(2), 119–137. https://doi.org/10.1093/bjc/azm073

Productivity Commission. (2024). *Report on government services: Justice*. Australian Government.

Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: Pathways and turning points through life*. Harvard University Press.

Sampson, R. J., & Laub, J. H. (2016). Turning points and the future of life-course criminology. *Journal of Research in Crime and Delinquency*, 53(3), 321–335. https://doi.org/10.1177/0022427815616992

Schwartz, M., Brown, D., & Cunneen, C. (2017). *Justice reinvestment* (Research Brief No. 21). Indigenous Justice Clearinghouse.

Scraton, P. (2007). *Power, conflict and criminalisation*. Routledge.

Smith, O., & Skinner, T. (2012). Observing court responses to victims of rape and sexual assault. *Feminist Criminology*, 7(4), 298–326. https://doi.org/10.1177/1557085112437875

Strang, H., Sherman, L. W., Mayo-Wilson, E., Woods, D., & Ariel, B. (2013). Restorative justice conferencing (RJC) using face-to-face meetings of offenders and victims. *Campbell Systematic Reviews*, 9(1), 1–59. https://doi.org/10.4073/csr.2013.12

Taylor, I., Walton, P., & Young, J. (1973). *The new criminology: For a social theory of deviance*. Routledge.

Taylor, P. J. (2014). Cross-cultural deception detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. 175–201). Wiley.

Telegraph, UK (2026), Retrieved from: https://www.telegraph.co.uk/multimedia/archive/03510/45751628_Australia_3510264b.jpg

Ugelvik, T., & Dullum, J. (Eds.). (2012). *Penal exceptionalism? Nordic prison policy and practice*. Routledge.

van der Kolk, B. (2014). *The body keeps the score: Brain, mind, and body in the healing of trauma*. Viking.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Wiley.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, 1(2), 110–117. https://doi.org/10.1016/j.jarmac.2012.02.004

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, 11(3), 89–121. https://doi.org/10.1177/1529100610390861

Walton, D., Martin, S., & Li, J. (2020). Iwi community justice panels reduce harm from re-offending. *Kotuitui: New Zealand Journal of Social Sciences Online*, 15(1), 75–92. https://doi.org/10.1080/1177083X.2019.1642921

Wexler, D. B., & Winick, B. J. (1996). *Law in a therapeutic key: Developments in therapeutic jurisprudence*. Carolina Academic Press.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

— End of Thesis —



Omxus Research, Perth WA 2026