

From biomedical principlism to policing ethics

NJAB+L and an elastic hierarchy of principles

Talk focus

Why we might think Principlism could work for Policing
Why the medicine → policing transfer needs re-engineering,
not a straight transplant. Then: what the adapted model adds.

Today's destination

The outline and defense of a practical framework
for making, documenting, and contesting ethical
trade-offs in policing.

Dr Hannah Maslen, Uehiro Oxford Institute for Practical Ethics (with credit to Colin Paine & Dr Matthew Minehan)

University of Melbourne | Health, Ethics, and Law Network (HELN) Seminar

3rd March 2026

Why principlism, and why policing ethics?

The problem: high-stakes value conflicts with weak action-guidance

- Police decisions routinely trade off values such as safety, rights, fairness, welfare, and trust
- Existing policing ethics tools often list values, but rarely tell you how to resolve conflicts
- Bioethics principlism is attractive because it combines a shared vocabulary with a method: specification + balancing
- But policing is not medicine, so the framework has to change in ethically principled ways

Aim of the talk

Explain the domain differences that necessitate adaptation, then sketch the adapted model and what it gets us.

Why medicine ≠ policing (ethically speaking)

Six domain differences that change the context for the principles

Medicine (typical case)

- Consent-forward relationships
- Therapeutic orientation to the patient
- More scope for deliberation and discussion
- Non-maleficence usually patient-directed
- Single person's interests at stake
- Legitimacy not a salient consideration (in medicine expertise is relevant and necessary)

Policing (typical case)

- Lawful coercion authorised by public consent – rights
- Adversarial and enforcement roles are routine
- Fast-time decisions under uncertainty
- Non-maleficence is multi-directional (public, suspect, victim, officer) – there is simply more going on
- Multiple people's interests at stake – balancing individual and collective goods
- Legitimacy is first-order democratic and instrumental value, grounded in public consent to coercive powers (and fragile!)

What changes when you move a framework across domains?

How the six differences reshape familiar principles

Autonomy

From self-determination (patient choice)
→ constraints on state power, rights, and agency.
Often relevant even when consent is absent and coercion justified

Beneficence

From acting for *this* patient
→ promoting **longer term** public welfare and security, including crime prevention.

Non-maleficence

From “do no harm” (mostly patient-focused)
→ multi-directional harm prevention across suspects, victims, bystanders, officers. **Immediate** harm from action and inaction

Justice

Criminal justice interests, distributive and procedural fairness

Legitimacy

Not optional. It is a boundary condition on coercive power, and a value with instrumental and democratic dimensions.

Our proposal: NJAB+L

A principlist vocabulary designed for policing

N Non-maleficence

J Justice

A Autonomy

B Beneficence

L Legitimacy (special role)

Method stays the same

Principlism as a *method*: specify what each principle means here, then balance when they conflict.

Content changes

Principles are specified for a coercive, publicly authorised domain, with explicit attention to legitimacy and multi-party harms.

Goal

Not perfect answers. Better reasons, clearer trade-offs, more transparent justifications.

Legitimacy is “special” (and why that matters)

Three roles: gate, constitution, and remainder

Stage 1: Bounded rightful authority (legitimacy gate)

If the proposed use of police power fails here, it is not eligible for balancing.

NJAB-constituted legitimacy

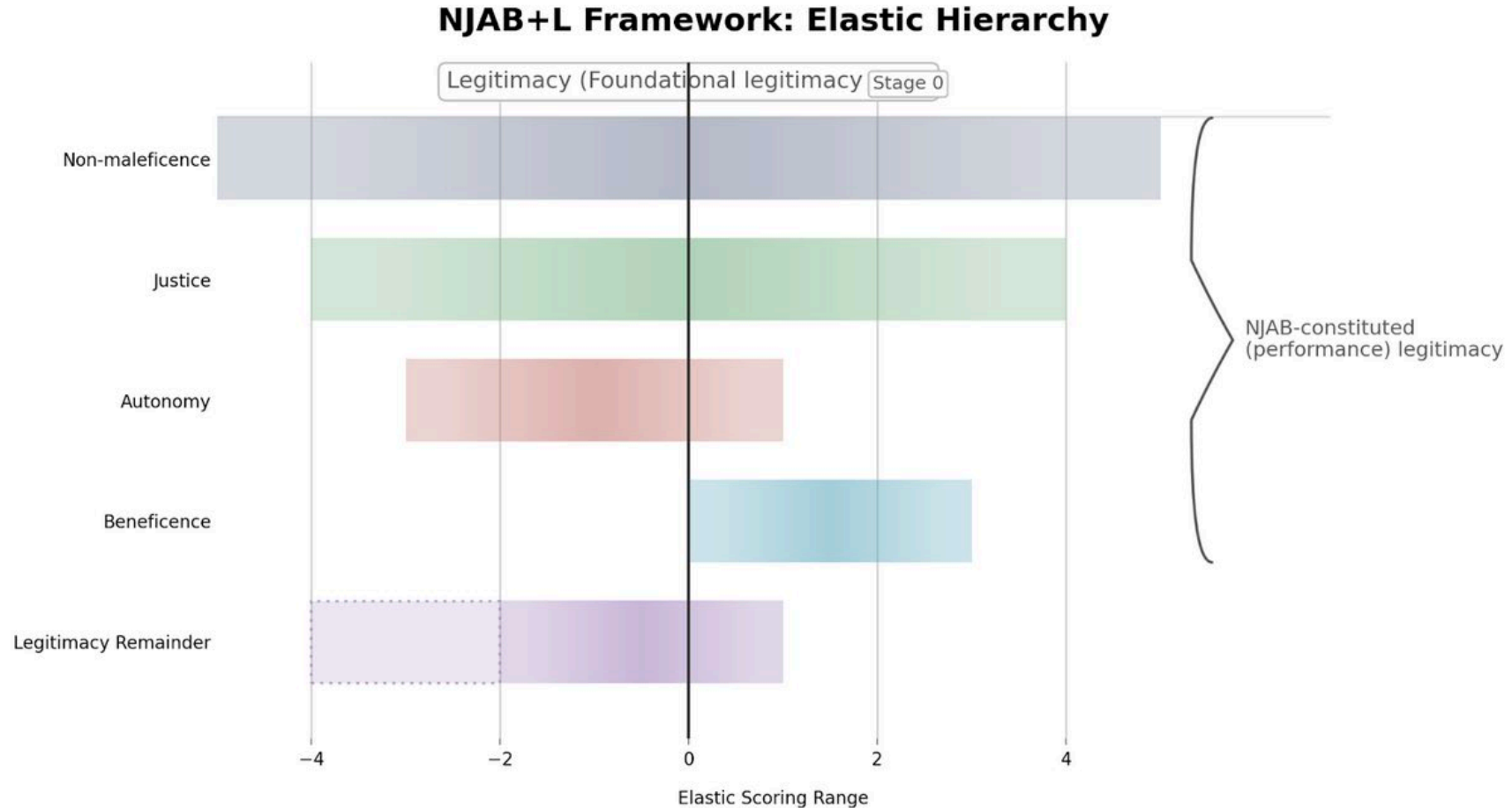
A lot of what makes policing legitimate is already captured by doing well on N, J, A, and B.
So we should not “double count” it under L as well.

Legitimacy remainder (scored separately)

- Instrumental: cooperation, compliance, trust
- Democratic: public standing to shape policing “in their name”
- Only what is not already captured by NJAB

An “elastic” hierarchy (not a rigid ranking)

Default weights, but they stretch with severity, probability, and urgency



Default ordering (after legitimacy gate): Non-maleficence > Justice > Autonomy > Beneficence > Legitimacy (remainder).
Ranges are illustrative and elastic.

How you actually use NJAB+L

A staged procedure for specifying, scoring, balancing, and revising

1 Legitimacy gate



2 Specify NJAB



3 Score NJAB



4 Score L remainder



5 Balance + justify



Two guardrails that matter in practice

- 1) Active-voice framing: “Is it ethically justifiable to exercise police power by doing X?”
- 2) Avoid double counting: define principles as mutually exclusive domains.

Scoring is a disciplined convention

- Use calibrated ranges (e.g., -5 to +5)
- Purpose: transparency and comparability, not “moral maths” – BUT we have argued for default ordering
- Numbers are revisable when you learn more or disagree

Output: a contestable justification

- What principle-domains were engaged?
- How strong were they in this case?
- Why did some outweigh others?
- What would change your mind?

Worked example 1: victim-opposed arrest

Criminal damage after an argument: should officers arrest?

Scenario (compressed)

A call-out after an argument. Ms C has caused **significant property damage** to vehicle, no injuries. The harmed party (Mr B) **does not want arrest** as it will upset their children, and Ms C has agreed to stay with a friend to cool down.

Active-voice question

“Is it ethically justifiable to **exercise police power by arresting** in this scenario?”

Indicative scoring (elaborated in the paper)

N: 0 J: 0 A: -2 B: +1 L(remainder): -1 → Total: -2

- Autonomy weighs heavily: arrest would disregard the expressed wishes of the harmed party
- Beneficence lightly supports intervention, but proportional alternatives exist
- Legitimacy remainder is negative: avoid escalating a low-risk incident in a way that undermines trust
- Conclusion: do not arrest; record, safeguard, signpost, and pursue proportionate alternatives

When might the outcome be different?

- Threats and risks of violence
- If the damage happened in a public not private setting

Principle	Ethically Relevant Considerations	Indicative Score
Non-maleficence	For Arrest (+1): A minor risk of escalation if the dispute continues, and a general interest in preventing further damage. Against Arrest (-1): Arrest is likely to be emotionally distressing and destabilising, including for the children if they wake. This is not severe harm, but it is a foreseeable, proximate cost.	0 (Net)
Justice	For Arrest (+1): A criminal offence has been committed, and the damage is significant enough to engage the criminal law. Against Arrest (-1): The criminal justice gain is limited. The victim is opposed, the case is likely to be resolved privately, and formal prosecution may be disproportionate relative to the nature of the wrongdoing and the availability of restorative options.	0 (Net)
Autonomy	Against Arrest (-2): Mr A is a competent adult explicitly refusing police intervention. Although he cannot “veto” enforcement, his wishes carry some weight, particularly given the private nature of the dispute and the absence of ongoing threat. Additionally, arrest would intrude significantly into family life late at night.	-2
Beneficence	For Arrest (+1): A modest deterrent effect. Arrest signals that property damage within intimate relationships is not exempt from law. However, the broader public welfare gains are limited.	+1
Legitimacy	Against Arrest (-1): A victim-opposed arrest in a low-risk, non-violent situation risks appearing heavy-handed and unnecessarily disruptive. This may weaken trust and future cooperation, especially if the public perceives policing as insensitive to proportionality and family stability.	-1

Worked example 2: Non-recent child sexual abuse (CSA)

Justice and prevention vs resource and harm to victim costs

Scenario (compressed)

During a review of 1990s safeguarding files from a summer holiday club, detectives find an overlooked note: **Girl F (18)** alleged volunteer **Mr G** “touched her under swimsuit”, with **no police referral**. **Ms F is now 38**, with complex PTSD and a history of self-harm; she has not come forward. **Mr G is now 80**, no record, and intelligence suggests **no regular contact with children**. Investigating would require trawling extensive archives and then **approaching Ms F unannounced**, risking severe re-traumatisation and consuming scarce specialist capacity.

Active-voice question

“Is it ethically justifiable to exercise police power by **initiating a non-recent CSA investigation** in this case (including contacting Ms F uninvited)?”

Indicative scoring (elaborated in the paper)

N: -1 J: +2 A: -1 B: +1 L(remainder): 0 → Total: +1

- Non-maleficence is a substantial constraint: unsolicited contact risks serious psychological harm to a known vulnerable survivor
- Justice is weighty: serious wrongdoing + rectifying historical injustice, but there are distributive justice costs (specialist time)
- Autonomy is mixed: acknowledgement and agency can be empowering, but uninvited approach is an intrusion; proceed only with sensitive options
- Beneficence supports investigating as a public good and possible harm-prevention, provided safeguards mitigate predictable harms
- Legitimacy remainder is neutral: public expects serious crimes to be pursued, but also expects ‘do no harm’ and proportionality

When might the outcome be different?

- Low solvability/high resource cost
- Active high risk access

Principle	Ethically Relevant Considerations	Indicative Score
Non-maleficence	<p>For Initiating (+1): There is no indication that Mr G has any regular current contact with children, either through volunteering or family. However, this cannot be absolutely precluded without investigation.</p> <p>Against Initiating (-2): There is a high risk of causing severe psychological harm to a known, vulnerable victim who has chosen not to contact the police herself. An unsolicited approach regarding this trauma could trigger a significant mental health crisis. This is a severe negative weight, albeit it could be mitigated with a multiagency approach and provision of specialist ongoing support.</p>	-1 (Net)
Justice	<p>For Initiating (+3): There is a strong <i>Criminal Justice</i> interest in investigating a serious historic crime (Child Sexual Abuse), compounded by the need to rectify a <i>historical injustice</i>.</p> <p>Against Initiating (-1): There is a <i>Distributive Justice</i> cost. Before approaching the victim, detectives will need to go back through thousands of old paper records related to this case. This will consume significant specialist resources, with an opportunity cost for other live and active cases.</p>	+2 (Net)
Autonomy	<p>For Initiating (+1): A sensitive approach could be seen as empowering, finally acknowledging the wrong done to Ms F and offering her the agency to seek justice.</p> <p>Against Initiating (-2): Approaching her uninvited is a significant intrusion into her privacy, overriding her silence of 30 years. The silence could be interpreted as an expression of a desire not to see a police investigation, albeit that cannot be confirmed without a visit. A visit is an <i>inquiry</i>, not a coercive act, so the score is -2 rather than -3.</p>	-1 (Net)
Beneficence	<p>For Initiating (+1): A successful investigation would be a public good, reinforcing societal condemnation of CSA and potentially having a small general deterrent effect.</p>	+1
Legitimacy	<p>For Initiating (+1): The public expects police to investigate serious crimes such as CSA and seek to bring the perpetrators to justice.</p> <p>Against Initiating (-1): The public also expects police to "Do No Harm." If the police are seen to cause the breakdown of a vulnerable survivor in pursuit of a speculative case, it risks damaging institutional trust.</p>	0 (Net)

What NJAB+L contributes (and what it does not)

A disciplined structure for justification, not a replacement for judgement

Contributions

- Shared vocabulary with clear domains
- Transparent trade-offs that can be audited
- Makes legitimacy explicit without letting it eclipse everything else
- Supports training, supervision, and public explanation

What it does NOT claim

- No “ethical calculator”
- Does not settle empirical disputes
- Does not eliminate reasonable disagreement
- Needs local specification and governance

Two likely objections (and short replies)

- 1) “Isn’t scoring reductive?” → It is a transparency tool; revise and contest the numbers.
- 2) “Why not virtue/care/rights-only?” → Those views still need a method for conflict cases; NJAB+L is compatible as a decision scaffold.

Thank you 😊

Happy to take questions or hear further objections!