Behavioural Insights in Criminal Law: Global Trends, Normative Challenges, and Emerging Frontiers

(Silvia Zorzetto)

Abstract

The intersection between behavioural sciences, cognitive psychology, and criminal law has been the subject of a vibrant and steadily expanding debate over the last decades. At the core of this discourse lies the attempt to integrate insights from behavioural law and economics (BLE) into penal theory and policymaking, with the goal of improving both the explanatory power of legal norms and the effectiveness of enforcement strategies. The initial waves of scholarship, often inspired by the heuristics-and-biases programme, sought to replace the idealised assumptions of rational choice theory with empirically grounded accounts of human decision-making. In the criminal law context, this substitution promised to refine traditional deterrence-based models, by tailoring legal rules and sanctions to actual cognitive patterns and motivational triggers.

The global debate has oscillated between enthusiasm and caution. On the one hand, proponents of BLE emphasise its capacity to predict behaviour more accurately than traditional rational actor models, and to design interventions – including nudges – that guide individuals toward socially desirable conduct without resorting to coercion. On the other hand, critical voices highlight several methodological objections. At the principled level, critics fear that criminal nudging policies risk treating individuals as mere objects of behavioural engineering, undermining the liberal conception of the offender as an autonomous moral agent. Moreover, the indeterminacy and context-sensitivity of behavioural predictions raise doubts about their reliability in the high-stakes domain of penal law, where errors have serious implications for both justice and legitimacy. Closely linked to this is the concern for the guarantees of the due process of law – namely, that criminal law must remain anchored in strict legality, proportionality, and fundamental rights – together with the underlying theories of punishment and crime.

The tension is compounded by normative challenges. In the theoretical-philosophical debate, the question arises whether and how legislators may legitimately interfere with the decision-making processes of potential offenders. There is also significant controversy regarding the type of ethical and meta-ethical assumptions, if any, embedded in the dimension of BLE and nudges in criminal law: whether they can be justified as deterrence-enhancing techniques, or whether they surreptitiously import paternalistic or consequentialist commitments that are in tension with liberal criminal law theory. These questions highlight the dual dimension of BLE in criminal law: it is not merely a descriptive science of how individuals actually behave, but also a normative enterprise that risks reshaping the foundations of punishment and deterrence. It is within this duality that many scholars have urged methodological caution, advocating for "interdisciplinary minimalism" – an approach that fosters cooperation across disciplines while avoiding the excesses of disciplinary imperialism.

Despite these concerns, the global diffusion of behavioural insights has only accelerated. From North America to Europe, Latin America, and Asia, the relevance of nudges and cognitive psychology for criminal law has grown more, not less, pronounced. Today, the debate has acquired even greater urgency for at least three reasons. First, the expansion of penal policy into areas such as cybercrime, financial misconduct, and transnational fraud requires an increasingly nuanced understanding of how offenders perceive risks, rewards, and normative constraints. Second, the rapid digitalisation of society has given rise to new forms of behavioural manipulation – often labelled as "dark patterns" – which both exploit cognitive biases and generate fresh challenges for criminal regulation. Third, the persistent ineffectiveness and inefficiency of traditional punitive systems increasingly casts doubt on their merely symbolic value – a hollow simulacrum – at a time when the demand for legality and effective protection of fundamental legal goods grows stronger. This creates a pressing challenge to devise innovative ways of pursuing effectiveness in criminal law remedies.

These most recent frontiers deserve closer scrutiny. In the paper, I shall focus exclusively on two: dark patterns and restorative justice. Dark patterns – deceptive digital architectures designed to manipulate users into making unintended choices – exemplify the double-edged nature of behavioural science in the penal domain. On the one hand, their existence confirms the predictive and manipulative power of cognitive insights: behavioural tendencies such as default bias, framing effects, and scarcity perception can be operationalised with striking effectiveness in the online environment. On the other hand, these very mechanisms, when deployed with exploitative intent, produce novel modalities of fraud, cybercrime, and money laundering. The challenge for criminal law is thus twofold: it must harness behavioural insights to design effective countermeasures, while at the same time grappling with the ethical and legal implications of confronting manipulation with

counter-manipulation. Regulators must also decide whether deceptive design techniques should be subsumed under existing categories of fraud and unfair commercial practices, or whether they require novel penal provisions. Given the global and borderless character of digital manipulation, jurisdiction, enforcement, and harmonisation emerge as critical issues. The penal response cannot rely solely on traditional doctrines of deterrence; rather, it must integrate empirical knowledge of cognitive vulnerabilities into both substantive criminal law and enforcement strategies.

Alongside these challenges, restorative justice represents an equally fertile and underexplored domain for BLE and nudges. Whereas most applications of behavioural sciences to criminal law have focused on deterrence, compliance, and enforcement, restorative justice places the emphasis on dialogue, accountability, and reintegration. This paradigm shift opens promising avenues for behavioural insights. Nudges could be designed not only to deter misconduct, but to foster empathy, to enhance perspective-taking, and to encourage voluntary engagement in restorative processes. Cognitive psychology suggests that offenders and victims are often locked in distorted perceptions, biases of attribution, and asymmetries of information that obstruct meaningful communication. Behavioural tools may help to overcome these obstacles by framing restorative encounters in ways that reduce defensiveness, promote trust, and make reparative gestures more salient. Moreover, restorative justice may itself serve as a test case for the normative legitimacy of behavioural interventions. By deploying nudges to strengthen voluntary participation and to support the internalisation of responsibility, behavioural sciences can show that their contribution need not be limited to paternalistic or manipulative techniques. Instead, they may help to design environments in which autonomy, dignity, and responsibility are enhanced rather than diminished. In this sense, restorative justice offers a normative counterweight to the scepticism that has often surrounded BLE in the penal sphere: it demonstrates how behavioural insights can be mobilised not for control, but for empowerment and reintegration.

In light of these developments, the balance of the debate shifts. Whereas earlier scholarship often oscillated between enthusiasm and scepticism about the feasibility of BLE in criminal law, today it seems increasingly clear that behavioural insights cannot be ignored. The challenge is not whether to integrate them, but how to do so responsibly, transparently, and in ways that enhance rather than erode the normative legitimacy of criminal law. The present contribution aims to provide both a stocktaking and a forward-looking perspective. It acknowledges the methodological and ethical criticisms that have been raised, and concedes that behavioural science cannot deliver a complete theory of criminal law. The task for scholars and policymakers is to ensure that the global diffusion of nudges and cognitive insights into penal discourse develops within clear boundaries, attentive to the risk of paternalism but equally aware of the opportunities for more humane and effective interventions.

Ultimately, the promise of behavioural sciences and nudges in criminal law lies not in supplanting traditional legal categories, but in enriching them. By integrating insights from psychology and behavioural economics, criminal law can better understand the actual drivers of compliance and deviance. By extending the scope of analysis to restorative justice and to the regulation of dark patterns, the field can also address emerging challenges that are reshaping the penal landscape worldwide. Far from being a marginal or passing debate, the behavioural turn in criminal law has become a central and inescapable dimension of global criminal policy.