TWO LOGICS, TWO SYSTEMS: THE EFFECTS OF INTERNATIONAL CRIMINAL JUSTICE THROUGH A BEHAVIORAL ECONOMIC LENS

Peter R. Grenzow[*]

ABSTRACT

A key question in the study of international criminal justice is what effects international prosecutions have on individuals' decisions to violate international criminal law (ICL). The literature has historically analyzed the effects of international prosecutions on individual decision-making through either the rationalist Logic of Expected Consequences (LEC) or the norms-based Logic of Appropriateness (LOA). However, such work has suffered from theoretical underdevelopment, with scholars largely failing to account for the insights of behavioral economics in their formulations of decision-making in response to prosecutions. This article fills this gap by analyzing the effects of international prosecutions with reference to the two systems of human cognition developed by behavioral economics: the unconscious System 1, with its various heuristics and biases, and the conscious System 2, underlying rational judgment. The article formalizes individuals' rational judgments in response to international prosecutions and discusses how Bayesian learning theory and

[*] International Humanitarian Law Instructor, American Red Cross. MPhil International Relations, Balliol College, University of Oxford. Juris Doctor, University of Minnesota Law School.

prospect theory may modify the assumptions of the rational-choice model. The article then discusses how the heuristics and biases of individuals' System 1 alter the rational judgments of System 2 in many situations. While this article provides a more refined account of individual judgment and decision-making in response to international prosecutions, it invites pessimism concerning the ability to develop generalizations about the consequences of international prosecutions. Considering these findings, prosecutors should analyze the complex and idiosyncratic psychology of the targets of their prosecutions to better understand how individuals may react to their legal actions and promote ICL's purpose of preventing crime.

TABLE OF CONTENTS

INTRODUCTION

> "Remember this punishment is for the purpose of prevention and not for vengeance."[1] – U.S. Secretary of War Henry L. Stimson

The deterrence of future crimes has been one of the principal purposes of both domestic criminal law[2] and international criminal law (ICL).[3] Through deterrence, criminal law enforcement seeks to disincentivize and thereby prevent the commission of crimes.[4] For instance, the Appeals Chamber of the International Criminal Tribunal for the former Yugoslavia (ICTY) recognized in the *Tadić* case that the "principle of deterrence …is a consideration that may

---

[1] GARY J. BASS, STAY THE HAND OF VENGEANCE: THE POLITICS OF WAR CRIMES TRIBUNALS 157 (2002) (quoting *McCloy-Stimson Conversation, 28 August 1944*, *in* BRADLEY F. SMITH, THE AMERICAN ROAD TO NUREMBERG: THE DOCUMENTARY RECORD, 1944-1945 23 (1981)) (U.S. Secretary of War Henry L. Stimson discussing post-World War II war crimes prosecutions).

[2] Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 COLUM. L. REV. 1193, 1193–95 (1985) (discussing deterrence and prevention as the "major function of criminal law"); Payam Akhavan, *Justice in the Hague, Peace in the Former Yugoslavia? A Commentary on the United Nations War Crimes Tribunal*, 20 HUM. RTS. Q. 737, 746 (1998) ("According to classical theory, one of the primary functions of criminal law is the deterrence of future criminal behavior.").

[3] Akhavan, *supra* note 2, at 743 (discussing how one of the purposes of the International Criminal Tribunal for the former Yugoslavia was preventing future atrocities); David Wippman, *Atrocities, Deterrence, and the Limits of International Justice*, 23 FORDHAM INT'L L. J. 473, 473 (1999) (discussing deterrence of future atrocities as a principal purpose of international criminal prosecutions).

[4] Akhavan, *supra* note 2, at 746.

legitimately be considered in sentencing…."[5] Moreover, in preambular paragraph five to the Rome Statute of the International Criminal Court (ICC), the State Parties affirm that they are "[d]etermined to put an end to impunity for the perpetrators of these crimes and thus to contribute to the prevention of such crimes…."[6] Considering this purpose, many studies have analyzed the effects of international prosecutions on individual judgment and decision-making, to determine whether and under what circumstances prosecutions may deter ICL violations.[7] Among other mechanisms, these studies have identified prosecutorial deterrence,[8] extra-legal social deterrence,[9] socialization or the internalization of ICL norms,[10] and the escalation of crimes as the effects of

---

[5] Prosecutor v. Tadić, Case No. IT-94-1-A & IT-94-1-A *bis*, Judgement in Sentencing Appeals, ¶ 48 (Int'l Crim. Trib. for the Former Yugoslavia Jan. 26, 2000).

[6] Rome Statute of the International Criminal Court pmbl. ¶ 5, July 17, 1998, 2187 U.N.T.S. 90 [hereinafter Rome Statute] (emphasis omitted).

[7] *See, e.g.*, Akhavan, *supra* note 2; Payam Akhavan, *Beyond Impunity: Can International Criminal Justice Prevent Future Atrocities?*, 95 AM. J. INT'L L. 7 (2001); Wippman, *supra* note 3; Theodor Meron, *From Nuremberg to the Hague*, 149 MIL. L. REV. 107 (1995); Hyeran Jo & Beth A. Simmons, *Can the International Criminal Court Deter Atrocity?*, 70 INT'L ORG. 443 (2016); Hunjoon Kim & Kathryn Sikkink, *Explaining the Deterrence Effect of Human Rights Prosecutions for Transitional Countries*, 54 INT'L STUD. Q. 939 (2010); Michael Broache, *International Prosecutions and Atrocities in the Democratic Republic of the Congo: A Case Study of the FDLR*, 7 J. MIDDLE E. & AFR. 19 (2016).

[8] Jo & Simmons, *supra* note 7, at 444 ("*Prosecutorial deterrence* is a direct consequence of legal punishment: it holds when potential perpetrators reduce or avoid law-breaking for fear of being tried and officially punished.").

[9] *Id.* ("*Social deterrence* is a consequence of the broader social milieu in which actors operate: it occurs when potential perpetrators calculate the informal consequences of law-breaking.").

[10] Broache, *supra* note 7, at 22 (discussing the socialization mechanism).

international prosecutions on individual decision-making.[11] The theories explaining these mechanisms generally rely on the assumption that the potential targets of international prosecutions process information about such legal actions through either a rationalist Logic of Expected Consequences (LEC) or a normative Logic of Appropriateness (LOA).[12] The importance of this work cannot be understated, because the explanation of individual decision-making processes in response to international prosecutions potentially gives prosecutors the ability to evaluate when their actions may lead to the prevention of crimes.

---

[11] *See* Kim & Sikkink, *supra* note 7, at 944 (discussing the escalation mechanism); Michael Broache, *Irrelevance, Instigation and Prevention: The Mixed Effects of International Criminal Court Prosecutions on Atrocities in the CNDP/M23 Case*, 10 INT'L. J. TRANSITIONAL JUST. 388, 393 (2016) ("Escalation operates by generating perverse incentives for belligerents to employ atrocities to secure victory or otherwise enhance their relative power in order to avoid future legal sanctions."); Jo & Simmons, *supra* note 7, at 445 (discussing the escalation effect of international prosecutions).

[12] *See* Broache, *supra* note 7, at 21-22 (discussing how rational-choice theories explain the deterrence mechanism while the "logic of appropriateness" explains the socialization mechanism). The "logic of expected consequences" assumes "that human actors choose among alternatives by evaluating their likely consequences for personal or collective objectives." James G. March & Johan P. Olsen, *The Institutional Dynamics of International Political Orders*, 52 INT'L ORG. 943, 949 (1998) [hereinafter March & Olsen, *The Institutional Dynamics*]. In contrast, the "logic of appropriateness" assumes that individuals match their behavior with a certain normative standard, which may include the proscriptions of a law, because they view such normative standards as legitimate or appropriate. James G. March & Johan P. Olsen, *The Logic of Appropriateness*, *in* THE OXFORD HANDBOOK OF POLITICAL SCIENCE 478, 478 (Robert E. Goodin ed., 2011) [hereinafter March & Olsen, *The Logic of Appropriateness*].

Nonetheless, these studies have largely depended on insights from criminology, focusing on the effects of domestic law enforcement on crime.[13] This approach makes sense because the international context of individual reactions to international prosecutions is analogous to the domestic context of offender reactions to law enforcement.[14] The criminological literature has in turn drawn its concepts from microeconomics because offenders' responses to law enforcement mechanisms involve decision-making processes which are the hallmark of the microeconomic field.[15] Especially since the mid-20th century, these microeconomic theories have relied upon rational-choice or expected utility theory, which describes human judgment and decision-making in the form of expected utility functions.[16] While objective, expected utility theories are normative, "in that they logically demonstrate how self-

[13] *See* Kim & Sikkink, *supra* note 7, at 943 (discussing the criminological literature studying deterrence); Jo & Simmons, *supra* note 7, at 447 (citing criminological studies in support of the prosecutorial deterrence mechanism); Broache, *supra* note 11, at 390 (discussing the criminological deterrence literature).

[14] *See* Broache, *supra* note 11, at 390-92 (applying the domestic criminological mechanisms of deterrence and incapacitation to the context of international prosecutions).

[15] *See generally* Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968); Daniel S. Nagin, *Criminal Deterrence Research at the Outset of the Twenty-First Century*, 23 CRIM. & JUST. 1 (1998) [hereinafter Nagin, *Criminal Deterrence Research*]; Daniel S. Nagin, *Deterrence in the Twenty-First Century*, 42 CRIM. & JUST. 199 (2013) [hereinafter Nagin, *Deterrence in the Twenty-First Century*].

[16] Greg Pogarsky, Sean P. Roche & Justin T. Pickett, *Offender Decision-Making in Criminology: Contributions from Behavioral Economics*, 1 ANN. REV. CRIMINOLOGY 379, 380-82 (2018) (discussing the microeconomic theories of offender decision-making).

interested actors can maximize their well-being."[17] Observing that the strong assumptions of expected utility theory may not come to pass in many circumstances, several criminologists have critiqued the theory as unrealistic of actual judgments in the context of criminal decision-making.[18]

Additionally, over the last half-century, with the pioneering work of scholars such as Herbert A. Simon, Amos Tversky, Daniel Kahneman, and Richard Thaler, cognitive psychologists and economists started to experimentally demonstrate that individual decision-makers systematically violate the assumptions underpinning expected utility theory.[19] This intersection of psychology and economics gave birth to the field of behavioral economics.[20] In contrast to the expected utility theories of microeconomics, behavioral economics is not merely normative, but instead "aims to develop 'descriptive economic models that accurately portray human

---

[17] *Id.* at 381 (citing RICHARD H. THALER, MISBEHAVING: THE MAKING OF BEHAVIORAL ECONOMICS (2015)).

[18] Raymond Paternoster, *How Much Do We Really Know about Criminal Deterrence*, 100 J. CRIM. L. & CRIMINOLOGY 765, 772-73 (2010) (discussing early critiques of expected utility theory).

[19] *See, e.g.*, Herbert A. Simon, *A Behavioral Model of Rational Choice*, 69 Q. J. ECON. 99 (1955); Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision under Risk*, 47 ECONOMETRICA 263 (1979) [hereinafter Kahneman & Tversky, *Prospect Theory*]; Amos Tversky & Daniel Kahneman, *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, 5 J. RISK & UNCERTAINTY 297 (1992) [hereinafter Tversky & Kahneman, *Advances in Prospect Theory*]; THALER, *supra* note 17.

[20] Daniel Kahneman, *Maps of Bounded Rationality: Psychology for Behavioral Economics*, 93 AM. ECON. REV. 1449 (2003) (discussing the intellectual roots of behavioral economics).

behavior.'"[21] Thus, behavioral economics provides a more realistic account of human judgment than expected utility theory, through the development of dual-process decision-making theory, accounting for the effects of judgment heuristics and biases, and prospect theory.[22]

While the criminological literature has recently started to incorporate the insights of behavioral economics,[23] the literature studying the effects of international prosecutions on individual decision-making has yet to do so. The remainder of this article seeks to fill this lacuna. Applying the insights of behavioral economics to the study of international criminal justice provides a more accurate account of the effects of international prosecutions on individuals' decisions to adhere to ICL. Consequently, it gives prosecutors the ability to evaluate the potential effects of their international prosecutions, including whether they may lead to the deterrence of ICL violations, with more precision. It also suggests several policies which may enhance the preventive effects of international prosecutions, beyond merely increasing arrests and convictions.

Section II reviews the existing literature analyzing the effects of international prosecutions on individual judgment and decision-making. It discusses the mechanisms of prosecutorial and social deterrence, socialization, and escalation, and the empirical evidence that supports the

---

[21] Pogarsky, Roche & Pickett, *supra* note 16, at 383 (quoting THALER, *supra* note 17.

[22] *Id.* at 383-93 (discussing the theories of behavioral economics).

[23] *See generally id.*; Greg Pogarsky, Sean P. Roche & Justin T. Pickett, *Heuristics and Biases, Rational Choice, and Sanction Perceptions*, 55 CRIMINOLOGY 85 (2017).

presence of these mechanisms in situations subject to international prosecutions. It focuses on these mechanisms in particular because they involve individual decision-making processes and may thus benefit from the insights of behavioral economics. Section III then reviews the two primary logics that the literature identifies as explaining individual judgments in response to international prosecutions, the LEC and the LOA.[24] The article describes these logics in terms of the two systems of human cognition, the unconscious and emotional System 1 and the conscious and deliberative System 2, theorized in behavioral economics.[25]

Section IV applies the deliberative models of System 2 judgment to explain the decision-making of individuals reacting to international prosecutions. It formalizes judgments accounting for international prosecutions in accordance with expected utility theory. It then discusses how Bayesian learning theory and prospect theory may explain changes in these individual judgments. Section V discusses the various heuristics and biases of System 1 and how they may alter the System 2 judgments of individuals responding to international prosecutions. Section VI concludes and discusses several policy implications derived from the insights of this article.

---

[24] *See* March & Olsen, *The Institutional Dynamics*, *supra* note 12, at 949-52.

[25] *See* Kahneman, *supra* note 20, at 1451 (discussing System 1 and System 2).

I. A REVIEW OF THE LITERATURE ON THE EFFECTS OF INTERNATIONAL PROSECUTIONS

The empirical record demonstrates that international prosecutions have mixed effects on individuals' commission of ICL violations, both preventing and promoting crimes. Prosecutions may prevent crime by deterring would-be criminals, incapacitating offenders, or socializing individuals to obey ICL.[26] Unfortunately, international prosecutions may also have the perverse effect of causing individuals to escalate their ICL violations, or they may empower actors to continue fighting and committing crimes by rallying support for them.[27] This article will focus on the mechanisms that deal directly with individual decision-making processes, including deterrence, socialization, and escalation. Nonetheless, because these mechanisms all assume that individuals both have knowledge of international prosecutions and change their behavior based on that knowledge, if either of these assumptions is not satisfied, international prosecutions may not affect individual decision-making.

A. Prosecutorial Deterrence

Prosecutorial deterrence is one of the primary mechanisms through which international prosecutions prevent crime.[28] Criminal deterrence theories assume that perpetrators are

---

[26] Broache, *supra* note 7, at 21-22 (discussing the deterrence, incapacitation, and socialization mechanisms).
[27] Broache, *supra* note 11, at 393-94 (discussing the escalation and empowerment mechanisms).
[28] Jo & Simmons, *supra* note 7, at 444.

rational actors who perceive a positive utility to be gained from crimes if the expected benefits to be achieved outweigh the expected costs, which are a product of the probability, celerity or immediacy, and severity of punishment.[29] Prosecutorial deterrence may be broken down into specific and general deterrence.[30] Specific prosecutorial deterrence refers to the deterrent effect of international prosecutions on the individuals subject to prosecution.[31] In contrast, general prosecutorial deterrence "is aimed at the discouragement of potential criminal behavior in society at large" and results in individuals not subject to prosecution choosing not to commit crimes due to the costs associated with potential international prosecutions.[32] In general, international prosecutions may deter crime by increasing the expected costs of crimes to the point at which the expected utility of ICL deviance is outweighed by the expected utility of ICL adherence.[33] While international prosecutions may thus have a direct effect on actors' cost-benefit calculations, they may also indirectly deter

---

[29] *See* Kim & Sikkink, *supra* note 7, at 943 (discussing criminal deterrence theory and the probability and severity of punishment); Broache, *supra* note 7, at 21 ("These theories treat potential perpetrators as rational actors who commit crime if its expected benefits exceed its expected costs, which are defined primarily in terms of the probability, severity, and celerity (swiftness) of legal sanctions….").

[30] *See* Leslie Vinjamuri, *Deterrence, Democracy, and the Pursuit of International Justice*, 24 ETHICS & INT'L AFF. 191, 194 (2010); Akhavan, *supra* note 7, at 12.

[31] Akhavan, *supra* note 2, at 746.

[32] *Id.*

[33] *See* Broache, *supra* note 7, at 21.

ICL violations by stimulating an increase in domestic prosecutions, which in turn deter crimes.[34]

Several cross-national studies have found empirical evidence supporting the deterrent effect of international prosecutions. Sikkink and Walling (2007) show that human rights trials in Latin America were associated with an improvement in countries' human rights practices.[35] Similarly, in a global study Kim and Sikkink (2010) find that "the level of repression in countries with [human rights] prosecutions is significantly lower than in countries without prosecutions."[36] Sikkink and Kim (2013) find that "[p]rosecutions of human rights violations are also associated with decreased use of torture, even if they do not reach convictions or if they end in acquittal."[37] Meernik (2015) finds that, among states that had ratified the Rome Statute of the ICC, the greater level of domestic rule of law was associated with "better human rights records and experience [of] less internal, political violence."[38] He also finds that state commitment to the ICC is associated with an improvement in human rights protections and a decrease in human rights abuses and political violence.[39] Jo and Simmons (2016) find that Rome Statute ratification and ICC prosecutions have a significant negative impact on civilian

---

[34] Jo & Simmons, *supra* note 7, at 448-49.
[35] Kathryn Sikkink & Carrie Booth Walling, *The Impact of Human Rights Trials in Latin America*, 44 J. PEACE RSCH. 427, 437 (2007).
[36] Kim & Sikkink, *supra* note 7, at 951-52.
[37] Kathryn Sikkink & Hun Joon Kim, *The Justice Cascade: The Origins and Effectiveness of Prosecutions of Human Rights Violations*, 9 ANN. REV. L. & SOC. SCI. 269, 282 (2013).
[38] James Meernik, *The International Criminal Court and the Deterrence of Human Rights Atrocities*, 17 CIV. WARS 318, 331 (2015).
[39] *Id.* at 333.

killing in situations under review of the court.[40] They also find evidence supporting an indirect deterrent effect of ICC prosecutions, showing that improvements in domestic criminal statutes, in reaction to ratification of the Rome Statute, had a negative impact on civilian killing.[41]

Case studies have bolstered these findings, explaining the causal mechanisms underlying prosecutorial deterrence. First, case studies demonstrate that combatants fear apprehension due to international prosecutions, and that fear may in turn cause a decrease in crime. In a case study of the *Congrés national pour la défense du peuple/Mouvement du 23-mars* (CNDP/M23) in the Democratic Republic of the Congo (DRC), Broache (2016) reports that Bosco Ntaganda's surrender to the ICC negatively impacted morale among the remaining M23 combatants, because they feared that they may be arrested for war crimes next.[42] Additionally, in Bosnia in the 1990s, camp commanders improved the circumstances of detainees in reaction to information about international prosecutions before the ICTY, suggesting that they feared potential prosecution for their maltreatment of detainees.[43]

Some empirical evidence also supports the specific prosecutorial deterrence mechanism. In a case study of the situation in Sudan, Wegner (2015) finds that in 2009, when it was public knowledge that the ICC Office of the Prosecutor

---

[40] Jo & Simmons, *supra* note 7, at 461.
[41] *Id.* at 462.
[42] Broache, *supra* note 11, at 406 ("[T]here is also some evidence suggesting that the negative effects of Ntaganda's surrender on M23's capacity resulted from fear of future legal sanctions among some combatants and supporters, consistent with the deterrence mechanism.").
[43] Akhavan, *supra* note 2, at 750-51.

(OTP) was pursuing a warrant against President Omar al-Bashir for crimes committed in Darfur, the OTP's actions may have contributed to the Government of Sudan's (GoS) entering into the Doha peace negotiations with the Justice and Equality Movement (JEM), to avoid the enforcement of an imminent warrant from the ICC.[44] Wegner also notes that "[a] high ranking JEM commander claimed that many foreign Janjaweed had left Darfur again due to fear of the ICC, and the UNAMID chief claimed that attacks on peacekeepers had diminished due to new rules of engagement and the threat of ICC prosecutions."[45] In a case study of the Lord's Resistance Army (LRA), Wegner finds that "the motivation to dodge an ICC warrant has been confirmed as the reason for the defection of an LRA commander" and the LRA "officers interviewed were clearly still worried about being indicted by the ICC today, even after defecting."[46] Wegner concludes that, at least initially between late 2004 and 2008, "the LRA refrained from large-scale massacres in the hope of not being indicted by the ICC."[47] While not conclusive evidence of deterrence, these studies support the functioning of the general and specific deterrence mechanisms of international prosecutions by stoking fear of prosecutions in the potential targets of such legal actions and raising the expected costs of committing crimes.

Nonetheless, case studies also demonstrate that international prosecutions have lacked a deterrent effect. For

---

[44] PATRICK WEGNER, THE INTERNATIONAL CRIMINAL COURT IN ONGOING INTRASTATE CONFLICTS: NAVIGATING THE PEACE-JUSTICE DIVIDE 106 (2015).
[45] *Id.* at 120 (citations omitted).
[46] *Id.* at 216 (citations omitted).
[47] *Id.* at 230.

instance, the Srebrenica massacre of more than 8,000 Bosniak Muslim men and boys was perpetrated by the Serbs in July 1995, after both the establishment of the ICTY and the indictment of Radovan Karadžić and Ratko Mladić, who were largely responsible for the atrocity.[48] Additionally, Slobodan Milošević continued to commit crimes in Kosovo after he was indicted by the ICTY.[49] The prosecutions of the International Criminal Tribunal for Rwanda (ICTR) and Rwanda's national courts did nothing to stop the Hutu *génocidaires*, the *Interahamwe* and later *Forces démocratiques de libération du Rwanda* (FDLR) from committing innumerable atrocities in Rwanda and the DRC.[50] Broache (2016) finds that ICC actions did not generate fear within the FDLR, and a FDLR "officer explicitly stated that learning about ICC indictments and arrests of leaders of other groups had not caused the FDLR to alter its strategy or tactics."[51] Similarly, Broache (2016) notes that the ICC warrant for Ntaganda, then-commander of the CNDP, "failed to trigger perceived increases in the probability, certainty or celerity of future legal action," required to deter the CNDP/M23 from committing atrocities.[52] This may have been due to the fact that the "CNDP maintained control over

---

[48] Akhavan, *supra* note 2, at 750; Wippman, *supra* note 3, at 480.
[49] Kim & Sikkink, *supra* note 7, at 943.
[50] *See* Wippman, *supra* note 3, at 482 ("[S]ome of the *genocidaires* responsible for mass murder in 1994 nonetheless continue to perpetrate atrocities in Rwanda from staging areas in the Congo."); Payam Akhavan, *Are International Criminal Tribunals a Disincentive to Peace?: Reconciling Judicial Romanticism with Political Realism*, 31 HUM. RTS. Q. 624, 626 (2009) (discussing the continued atrocities of the *génocidaires* in the DRC).
[51] Broache, *supra* note 7, at 30-31.
[52] Broache, *supra* note 11, at 401.

large swaths of territory from which it effectively excluded government forces" and "Ntaganda personally enjoyed a relatively protected position…."[53] Wegner (2015) finds that, in Darfur, the perceived probability of punishment and the knowledge of the ICC among the rank-and-file soldiers of the JEM and Janjaweed was low, and these factors may have contributed to the failure of ICC prosecutions of rebel leaders to deter attacks on peacekeepers in the region.[54] Further, while ICC actions may have initially deterred the LRA, once the group failed to avoid the ICC warrants through the Juba Talks in 2008 it resumed committing atrocities, and the ICC failed to have a further deterrent effect.[55] Kersten (2016) also notes the lack of a deterrent effect of the ICC in Libya, evidenced by the fact that Saif al-Islam Gaddafi, in discussing the ICC's warrant against him and his father, viewed the ICC as "a joke" and "controlled by those countries which are attacking us every day!"[56] Thus, the existing literature suggests that international prosecutions have had mixed results in deterring ICL violations. Aside from deterrence stemming from fear of formal legal sanctions, international prosecutions may deter crimes by causing combatants to consider the social costs of crimes.

---

[53] *Id.*
[54] WEGNER, *supra* note 44, at 124-25.
[55] *Id.* at 230.
[56] MARK KERSTEN, JUSTICE IN CONFLICT: THE EFFECTS OF THE INTERNATIONAL CRIMINAL COURT'S INTERVENTIONS ON ENDING WARS AND BUILDING PEACE 128 (2016).

B. Social Deterrence

Along with prosecutorial deterrence, social deterrence is another mechanism through which international prosecutions may prevent crime. Social deterrence pertains to individuals' consideration of the informal, extralegal consequences of charge, apprehension and sanction resulting from international prosecutions, including the costs of stigmatization, delegitimization, losing material support, economic sanctions, and naming-and-shaming.[57] Like prosecutorial deterrence, social deterrence theory assumes that individuals are rational actors and will choose not to commit ICL violations if the expected extralegal costs resulting from such crimes sufficiently outweigh the benefits to be gained.[58]

Jo and Simmons (2016) find some evidence supporting social deterrence in their global study. They show that, among states which ratified the Rome Statute, an increase in human rights organizations present—indicating potential naming-and-shaming—was associated with a decrease in intentional civilian killing.[59] Additionally, an increase in aid received among states that had ratified the Rome Statute—indicating potential social pressure resulting from the possibility of losing aid—was also negatively associated with civilian killing.[60] Nevertheless, case studies are needed to confirm the causal

---

[57] *See* Jo & Simmons, *supra* note 7, at 444, 449-52 (defining social deterrence); Kim & Sikkink, *supra* note 7, at 945 ("Informal social sanctions may follow from the formal sanctions of trials and can have important effects in political arenas where reputation is essential.").
[58] Jo & Simmons, *supra* note 7, at 450.
[59] *Id.* at 463.
[60] *Id.* at 464.

mechanisms explaining social deterrence. In addition to being deterred from committing crimes, international prosecutions may also cause combatants to be socialized towards adherence to ICL.

C. Socialization

Moving beyond the rationalist theories of international prosecutions, Broache (2016)[61] elaborates upon the unconscious prevention mechanism identified by Akhavan (2001),[62] through which international prosecutions have a pedagogical influence on individuals' perceptions about what is acceptable behavior. Broache defines this mechanism as "socialization," through which international prosecutions prevent atrocities "by altering underlying 'logics of appropriateness'" by demonstrating that atrocities are not normatively acceptable.[63] Specifically, individuals internalize or habituate the normative standard providing that ICL violations are wrong, and, as a result, crimes do not present themselves as acceptable conduct in the future.[64] In its judgment in the *Kordić and Čerkez* case, the Appeals Chamber of the ICTY has also discussed the socialization mechanism as follows:

> One of the most important purposes of a sentence imposed by the International Tribunal is to make it

[61] Broache, *supra* note 7, at 22.
[62] *See* Akhavan, *supra* note 7, at 12-13.
[63] Broache, *supra* note 7, at 22 (quoting March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 689-708).
[64] *Id.*

> abundantly clear that the international legal system is implemented and enforced. This sentencing purpose refers to the educational function of a sentence and aims at conveying the message that rules of humanitarian international law have to be obeyed under all circumstances. In doing so, the sentence seeks to internalise these rules and the moral demands they are based on in the minds of the public.[65]

However, in his case study, Broache concludes that, among the FDLR members interviewed, ICC actions against leaders of other armed groups "failed to alter understanding of the appropriateness of atrocities, which is required for socialization."[66] This may have been due to the fact that the FDLR, composed of the ethnic Hutu group, dehumanizes the Tutsis in Rwanda and view killing the Tutsis as the morally correct standard of appropriateness.[67] In any event, more research is needed to trace the effect of international prosecutions in socializing individuals. Unfortunately, as the studies discussed above demonstrate, while international prosecutions may cause a decrease in crime, they may also lead to an escalation in atrocities.

---

[65] Prosecutor v. Kordić and Čerkez, Case No. IT-95-14/2-A, Judgement in Appeal, ¶ 1080 (Int'l Crim. Trib. for the Former Yugoslavia Dec. 17, 2004).
[66] Broache, *supra* note 7, at 31.
[67] *See id*. at 31-32.

D. Escalation

International prosecutions may have the perverse effect of promoting the escalation of ICL violations. Assuming that actors rationally weigh the expected costs and benefits of their conduct, the escalation mechanism provides that individuals perceive a positive utility to be gained by escalating their hostilities and associated atrocities in reaction to international prosecutions.[68] For example, instead of negotiating for peace, individuals may escalate their ICL violations to increase or maintain their power and avoid international prosecutions.[69] International prosecutions may generally promote the escalation of atrocities by those individuals not subject to prosecutions, or they may specifically instigate the escalation of atrocities by those offenders subject to legal actions.[70] However, specific escalation may function more strongly than general, because the offenders subject to international prosecutions may perceive a greater expected loss from not committing atrocities to avoid prosecution.[71] This may result from the fact that they perceive a greater expected probability of apprehension and sanction for their crimes, and thus

---

[68] *See* Kim & Sikkink, *supra* note 7, at 944; Broache, *supra* note 11, at 393; Jo & Simmons, *supra* note 7, at 445.
[69] Broache, *supra* note 11, at 393 ("[A]trocities may be useful, or at least perceived as such, for actors seeking to secure victory or otherwise enhance their relative power by eliminating opponents, deterring collaboration with enemies and/or building organizational cohesion.").
[70] *Id*.
[71] *Id*.

perceive a greater overall utility to be gained from escalating hostilities and atrocities to avoid prosecution.[72]

Case studies from several contexts subject to ICC prosecutions support the presence of the escalation mechanism. Broache (2016) found that the ICC warrant for Sylvestre Mudacumura of the FDLR may have caused "Mudacumura to spoil peace initiatives, plausibly contributing to the uptick in violence beginning in mid-2014."[73] Broache also found that the 2012 ICC conviction of Thomas Lubanga Dyilo and subsequent demands for the arrest of Bosco Ntaganda in the Eastern DRC caused Ntaganda to form the M23, in order to avoid arrest, which led to an increase in hostilities and atrocities in the region.[74] Scholars have further found that the ICC warrant for Omar al-Bashir prompted him to expel and harass several humanitarian nongovernmental organizations (NGO), which were providing approximately 40% of the humanitarian aid to internally displaced persons (IDP) in Darfur.[75] This in turn led to an increase in malnutrition and food shortages in the region.[76]

Additionally, the ICC warrants for the LRA leadership in 2005 correlated with a subsequent spike in the number of attacks by the LRA on staff members of international humanitarian organizations in the region.[77] On the same day that the attacks started, LRA commander Vincent Otti

---

[72] *See id*.
[73] Broache, *supra* note 7, at 34.
[74] Broache, *supra* note 11, at 404-05.
[75] Vinjamuri, *supra* note 30, at 196; WEGNER, *supra* note 44, at 133-34.
[76] WEGNER, *supra* note 44, at 134.
[77] *Id*. at 265.

confirmed that the LRA had knowledge of the ICC warrants, and "there were reports of LRA threats to kill Westerners as a reaction to the ICC indictments," suggesting the ICC warrants caused the escalation in attacks against the organizations.[78] Towards the end of 2008, after the LRA perceived that it would be unable to dodge the ICC warrants against its leadership through the Juba Talks, the LRA vastly increased its attacks on civilians in the DRC, southern Sudan and the Central African Republic (CAR).[79] In line with the escalation mechanism, the leadership of the LRA likely concluded that, because they would face a trial before the ICC if they surrendered, they would gain a higher utility from fighting to avoid apprehension and incarceration.[80] Thus, the empirical evidence supports the escalation effect of international prosecutions.

The studies surveyed in this section have done much to advance the understanding of the effects that international prosecutions have on the commission of ICL violations. However, past studies have failed to account for advances in behavioral economics and cognitive psychology. Applying behavioral economics to the study of international criminal justice is important because it offers a more precise and realistic understanding of how individuals may react to international prosecutions. Toward that end, the remainder of this article discusses how the two systems of human cognition,

---

[78] *Id*.
[79] *Id*. at 231; *see also The Christmas Massacres: LRA Attacks on Civilians in Northern Congo*,
HUM. RTS. WATCH (2009), https://www.hrw.org/sites/default/files/reports/drc0209webwcover_1.pdf (describing the LRA attacks in the DRC).
[80] WEGNER, *supra* note 44, at 231.

defined by behavioral economics, interact in producing the varying reactions of individuals to international prosecutions.

II. INDVIDUAL REACTIONS TO INTERNATIONAL PROSECUTIONS THROUGH A BEHAVIORAL ECONOMIC LENS

As introduced above, international prosecutions may affect individuals' decisions to adhere to International Criminal Law (ICL) through the rationalist Logic of Expected Consequences (LEC) and the norm-based Logic of Appropriateness (LOA). In turn, these two logics can both be explained in terms of the two systems of cognition and their respective psychological mechanisms. This section first defines the two logics and explains the mechanisms discussed in the previous section in terms of the logics and their interactions. This section then explains the two logics with regard to the two systems of human cognition drawn from behavioral economics.

A. Two Logics of Law-Influenced Conduct

There are two primary logics of action[81] which govern behavior in situations involving normative standards: the LEC and the LOA.[82] As a preliminary matter, the term "norm" is generally defined "as a standard of appropriate behavior for

[81] March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 485 ("Logics of action are used to describe, explain, justify, and criticize behavior….").
[82] March & Olsen, *The Institutional Dynamics*, *supra* note 12, at 949-52 (defining the logics); Broache, *supra* note 7, at 22 (discussing the LOA).

actors with a given identity…."[83] The norms literature has generally identified two types of norms, including: 1) injunctive norms, which "inform us about what is typically approved/disapproved," and 2) descriptive norms, which "inform us about what is typically done…."[84] This article focuses on the former type of norms, to which the proscriptions of ICL pertain.

The LEC assumes that individual behavior is driven by the expected consequences of actions, and "human actors choose among alternatives by evaluating their likely consequences for personal or collective objectives…."[85] The LEC also assumes that actors are rational, meaning that actors with the same information reason to the same conclusions concerning such information,[86] employing cost-benefit calculations to meet their objectives.[87] Under the LEC, actors will adhere to ICL if doing so produces a maximal expected utility.[88] The LEC thus corresponds to rational-choice theory, which forms the basis of the mechanisms of general and specific prosecutorial deterrence, social deterrence, and escalation, elaborated upon

---

[83] Martha Finnemore & Kathryn Sikkink, *International Norm Dynamics and Political Change*, 52 INT'L ORG. 887, 891 (1998); *see also* Vaughn P. Shannon, *Norms Are What States Make of Them: The Political Psychology of Norm Violation*, 44 INT'L STUD. Q. 293, 294 (2000).

[84] Robert B. Cialdini & Noah J. Goldstein, *Social Influence: Compliance and Conformity*, 55 ANN. REV. PSYCH. 591, 597 (2004); *see also* P. Wesley Schultz et al., *The Constructive, Destructive, and Reconstructive Power of Social Norms*, 18 PSYCH. SCI. 429, 430 (2007).

[85] March & Olsen, *The Institutional Dynamics*, *supra* note 12, at 949.

[86] James D. Fearon, *Rationalist Explanations for War*, 49 INT'L ORG. 379, 392 (1995).

[87] Shannon, *supra* note 83, at 295 ("This logic is individualist in orientation, where rational, cost-benefit calculations drive actors to meet their desired goals.").

[88] *See id*. at 295-96.

above.[89] As these mechanisms, and the cited evidence supporting them, suggest, the effect of international prosecutions on individuals' decisions to adhere to ICL is mediated in part through the LEC.

In contrast to the LEC, the LOA assumes that actors comply with ICL norms, not because doing so produces a maximal expected utility, but because the normative obligations are perceived as "natural, rightful, expected, and legitimate."[90] Individuals thus act in accordance with norms on the basis of how similar their action is to the prescribed normative standard, instead of considering the expected utility to be obtained from such action.[91] While "appropriateness" suggests some inherent moral "good," standards of appropriateness are inter-subjective understandings of "what is true, reasonable, natural, right, and good" and could thus form the basis of "atrocities of action, such as ethnic cleansing and blood feuds, as well as moral heroism."[92] Depending on how internalized a given normative standard is, the LOA may be processed through various levels of conscious and unconscious judgment.[93] Unconscious or habit-based action under the LOA explains the socialization mechanism,[94] discussed in the previous section. The studies outlined above thus suggest that

---

[89] *See supra* Section II.
[90] March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 478; *see also* Finnemore & Sikkink, *supra* note 83, at 912.
[91] March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 479.
[92] *Id*.
[93] *Id*.; *see also* Finnemore & Sikkink, *supra* note 83, at 905.
[94] Broache, *supra* note 87, at 22; *see also* Akhavan, *supra* note 37, at 13.

the effect of international prosecutions on individual decisions to adhere to ICL may also be mediated by the LOA.[95]

The LEC and the LOA are not mutually exclusive, and either or both logics may be involved in individual reactions to international prosecutions.[96] This is made clear by the literature reviewed above,[97] which discusses evidence supporting the presence of decision-making based on expected utility and on standards of appropriateness in contexts involving international prosecutions. One rule that may govern the implementation of one logic of action over the other provides that the logic that is more clearly applicable to a situation may dominate the other logic.[98] For instance, if knowledge among individuals of international prosecutions is low, the expected costs and benefits of committing an ICL violation may be ambiguous. In contrast, the individuals' normative standard may be clearly applicable to the situation, in which case the LOA will dominate the LEC.

Empirical evidence confirms that combatants' LOAs may have dominated their LECs with the result of preventing any deterrent effect of international prosecutions. For instance, in the case of the FDLR, limited knowledge of the ICC, coupled with a clear normative standard promoting atrocities against the Tutsis, may have prevented the deterrent impact of ICC actions.[99] Similarly, the normative standards of the Lord's Resistance Army (LRA), who "did not perceive themselves as

---

[95] *See supra* Section II.C.
[96] *See* March & Olsen, *The Institutional Dynamics*, *supra* note 12, at 952; Shannon, *supra* note 83, at 298.
[97] *See* Section II.
[98] March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 492.
[99] *See* Broache, *supra* note 7, at 29-32.

soldiers, but rather as teachers of God's message," may have prevented a deterrent effect of any prospective sanctions from the ICC.[100]

While the two logics of action thus provide a solid theoretical framework for describing individual reactions to international prosecutions, identifying the psychological mechanisms underlying the two logics allows for the more precise explanation of why individuals either adhere to or violate ICL in response to prosecutions. Moreover, accounting for these psychological mechanisms allows for the explanation of individual judgment that does not fall squarely within either of the two logics of action. The next section will explain the two logics in terms of the two systems of cognition.

B. Two Systems of Cognition

The LEC and the LOA may be explained by the two systems of cognition, drawn from the dual-process decision-making model of behavioral economics.[101] Dual-process

---

[100] Sarah Kihika Kasande, *Evaluating the Deterrent Effect of the International Criminal Court in Uganda*, *in* TWO STEPS FORWARD, ONE STEP BACK: THE DETERRENT EFFECT OF INTERNATIONAL CRIMINAL TRIBUNALS, *infra* note 124 at 201, 214.

[101] *See* Pogarsky, Roche & Pickett, *supra* note 16, at 388-90 (discussing dual-process decision-making theory and its application to criminology); DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011) (discussing the two systems of cognition); Robert Apel, *Sanctions, Perceptions, and Crime: Implications for Criminal Deterrence*, 29 J. QUANTITATIVE CRIMINOLOGY 67, 92 (2013) ("Of most relevance for criminologists is the development of dual-process models of decision making that integrate two cognitive systems—reasoned (analytical, deliberative, 'cold' state) and reactive (intuitive, affective, 'hot' state) decision making.").

decision-making theory provides that all cognitive judgments are governed by two systems, System 1 and System 2, which, like the two logics, may independently or interactively mediate decision-making.[102] System 1 is generally described as "intuitive, automatic, natural, non-verbal, narrative, and experiential…."[103] The system "provides constant and near instantaneous answers to the questions in daily life" including "questions about the risk of future events" and "is able to provide the answers through the use of mental shortcuts, also known as cognitive heuristics" and their attendant biases.[104] Thus, this system corresponds to humans' unconscious judgments. An example of System 1 decision-making in the criminal context is when a driver stops at a red light or a stop sign out of habit, without considering running the light or sign consciously.[105]

System 2 judgment is conscious, deliberate, and slower than that of System 1.[106] System 2 thus exemplifies the reasoning that is assumed in models of rational choice, Bayesian updating, and prospect theory, discussed *infra*.[107] System 2 judgment would be involved if a driver, when approaching a red light, considers the expected costs and benefits that would result from the driver running the red light,

---

[102] Kahneman, *supra* note 20, at 1450-52; Paul Slovic, *"If I Look at the Mass I will Never Act": Psychic Numbing and Genocide*, 2 JUDGMENT & DECISION MAKING 79, 84 (2007).
[103] Seymour Epstein, *Integration of the Cognitive and the Psychodynamic Unconscious*, 49 AM. PSYCH. 709, 710 (1994); *see also* Slovic, *supra* note 102, at 82; Kahneman, *supra* note 20, at 1451.
[104] Pogarsky, Roche & Pickett, *supra* note 16, at 388.
[105] *See generally* Kahneman, *supra* note 20, at 1450-52.
[106] Epstein, *supra* note 103, at 710; *see also* Slovic, *supra* note 102, at 82; Kahneman, *supra* note 20, at 1451.
[107] *See* Pogarsky, Roche & Pickett, *supra* note 16, at 388-89.

such as a traffic ticket or getting to work on time.[108] While most judgments are conducted through the effortless thought of System 1, System 2 may monitor the intuitive judgments formed by System 1, rejecting or modifying the heuristic answers provided by it.[109] However, "the monitoring is normally lax, and allows many intuitive judgments to be expressed, including some that are erroneous."[110] Thus, System 2 often merely ratifies or utilizes the intuitive judgments formed by System 1.[111] In this way, like the LEC and the LOA, the two systems interact in mediating decision-making.

The LOA may involve both systems of judgment. First, the matching of behavior with a standard of appropriateness may be based entirely on experience, habit and intuition, thus taking place within System 1.[112] Pairing conduct with a standard of appropriateness may also be "a relatively complicated cognitive process involving thoughtful, reasoning behavior," suggesting System 2 cognition.[113] Judgment in accordance with the LOA through System 1 may result for a number of reasons. System 1 may dominate System 2 in situations involving "time pressure," "concurrent involvement in a different cognitive task," "performing the task in the evening for 'morning people' and in the morning for 'evening people,'" and "being in a good mood."[114] System 2 judgment is effortful,

---

[108] *See id*.
[109] Kahneman, *supra* note 20, at 1450.
[110] *Id*. (citation omitted).
[111] Pogarsky, Roche & Pickett, *supra* note 16, at 388.
[112] March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 479.
[113] *Id*.
[114] Kahneman, *supra* note 20, at 1451 (citations omitted).

energy-intensive, and time-consuming.[115] If a combatant has to make a split-second decision, the individual may rely entirely on their intuitive judgments, basing their decision on an internalized normative standard.[116] If the combatant is afforded more time to make the judgment, the judgment may take place largely within System 2 and be more deliberative.[117] If a context evokes no new information regarding a standard of appropriateness or expected utilities, the individual may merely rely on habit and System 1 in acting in accordance with a norm.[118] In contrast, if the situation involves new information, the judgment may become more energy-intensive and consider the legitimacy of the potentially-applicable norms using System 2.[119]

In accordance with the availability heuristic, the choice between multiple norms may be the product of the ease of access to the normative standard in the individual's memory.[120] Accessibility of a norm may result from several psychological factors. "Personal, emotionally involving events are much more likely to be remembered than information that is less immediate or concrete."[121] This suggests that, if an ICL norm is associated with an emotionally involving event, such as a combatant's initial exposure to the prohibition on targeting

---

[115] *Id*.
[116] *See* March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 493.
[117] *See id*. at 479.
[118] Kahneman, *supra* note 20, at 1451.
[119] *See id*.
[120] *See id*. at 1452; Pogarsky, Roche & Pickett, *supra* note 23, at 93 (discussing the availability heuristic).
[121] ROSE MCDERMOTT, POLITICAL PSYCHOLOGY IN INTERNATIONAL RELATIONS 64 (2004) (citation omitted).

civilians in the middle of an armed engagement, it may be more available and more likely to come to mind. Additionally, norms will be more available if they "are familiar or salient, because they have been seen repeatedly or recently…."[122] Thus, if ICL norms are transmitted to combatants repeatedly and tied to emotionally compelling content—such as graphic imagery—they may become more psychologically available to the individuals in the future, promoting ICL adherence. The availability heuristic may therefore explain the pedagogical influence of international prosecutions, underlying the socialization mechanism,[123] through which prosecutions promote ICL adherence by educating individuals on ICL norms.

Suggestive evidence for this conclusion has been noted in the Eastern DRC, where community leaders reported that, following the ICC's actions concerning the context, "there now exists greater knowledge that recruitment and use of children is a violation of the law" and that "awareness has translated into lower numbers of children, particularly in self-defense militias."[124] Castano, Muñoz-Rojas, and Čehajić-Clancy (2020) also support this conclusion with experimental evidence, showing that greater knowledge of International

---

[122] *Id*. at 65 (citations omitted); *see also* March & Olsen, *The Logic of Appropriateness*, *supra* note 12, at 484.

[123] *See supra* Section II.C.

[124] Sharanjeet Parmar, *Dissuasive or Disappointing? Measuring the Deterrent Effect of the International Criminal Court in the Democratic Republic of the Congo*, *in* TWO STEPS FORWARD, ONE STEP BACK: THE DETERRENT EFFECT OF INTERNATIONAL CRIMINAL TRIBUNALS 173, 181 (Jennifer Schense & Linda Carter eds. 2016) (quotations and citation omitted).

Humanitarian Law (IHL) was associated with greater intentions to respect IHL among combatants.[125] Similarly, a recent survey experiment, conducted for this article, supports the same conclusion. In the experiment, involving 1,073 adults located in the United States, prior knowledge of the proscriptions of IHL, correlating with the norms of ICL, was associated with a lower acceptance of ICL violations among participants.[126] This evidence suggests that the prior knowledge of ICL norms among individuals, and thus the greater availability of the normative standards, makes individuals more likely to adhere to ICL in accordance with the availability heuristic and the socialization mechanism.

In contrast to the LOA, the LEC pertains primarily to the conscious, deliberative, effortful and reasoned judgments of System 2.[127] This system conducts the energy-intensive reasoning that is assumed in economic expected utility models, Bayesian updating models, and prospect theory, elaborated upon below.[128] However, the System 2 utility calculations of the LEC may be affected and skewed by the processes of

---

[125] Emanuele Castano, Daniel Muñoz-Rojas & Sabina Čehajić-Clancy, *Thou Shalt Not Kill: Social Psychological Processes and International Humanitarian Law Among Combatants*, 26 PEACE & CONFLICT: J. PEACE PSYCH. 35, 39 (2020); *see also* Geoffrey P.R. Wallace, *Martial Law? Military Experience, International Law, and Support for Torture*, 58 INT'L STUD. Q. 501, 507 (2014).

[126] Peter R. Grenzow, *2021 MTurk United States IHL Survey Code and Data*, HARV. DATAVERSE (Oct. 4, 2021), https://doi.org/10.7910/DVN/A8LABR, at Table 4.2 (showing that increases in the *IHLScore*, indicating subjects' knowledge of IHL, is associated with increases in the *MoralScore*, indicating lower acceptance of ICL violations).

[127] *See* March & Olsen, *The Institutional Dynamics*, *supra* note 12, at 949; Kahneman, *supra* note 20, at 1451.

[128] *See* Pogarsky, Roche & Pickett, *supra* note 16, at 388-89; Section IV.

System 1 and the LOA.[129] For instance, individuals considering an ICL violation may form intuitive impressions concerning the probability of apprehension, and the formal or informal sanctions for the crime, in their System 1 judgment. These intuitive impressions may then be rejected, affirmed, or modified by their System 2 thinking.[130] The next section will discuss in greater depth individuals' System 2 judgments, in reaction to international prosecutions.

III. SYSTEM 2 DELIBERATIVE MODELS OF JUDGMENT

Expected utility, Bayesian learning, and prospect theory are widely employed in criminology to explain offenders' deliberative System 2 decision-making. This section applies these theories to the analysis of international criminal justice to provide a more precise and realistic account of individual reactions to international prosecutions. It first formalizes how individuals may calculate their expected utilities concerning prospective ICL violations, then explains how individuals' judgments may change over time in accordance with Bayesian learning theory, and finally discusses how loss aversion may modify the deliberative judgments of some individuals.

A. Expected Utility Theory

Expected utility theory assumes that decision-makers are "goal-oriented and self-interested individuals who are

---

[129] *See infra* Section V.
[130] Pogarsky, Roche & Pickett, *supra* note 16, at 388.

endowed with stable and well ordered 'preferences,' and who judge different courses of action by their expected consequences … in such a way as to maximize their utility and overall satisfaction."[131] In judging the expected consequences of different prospective choices, individuals consider the potential costs, or losses, and potential benefits, or gains, of different courses of action.[132] However, because the future is governed by uncertainty, the potential costs and benefits considered are conditioned by the expected probability of their occurrence.[133] Under expected utility theory, actors are generally assumed to be risk averse.[134] This means that, if given the choice between a gamble involving the risk of loss and a sure-thing, the actor "would accept an amount less than the average monetary value of a gamble rather than take that gamble."[135]

Individual judgment is thus assumed to be governed by a utility function *U(*)*, in which the utility, or the monetary or psychological value of a choice, is a function of the expected losses and gains of the choice and the probabilities of incurring the same.[136] Drawing on Pogarsky, Roche, and Pickett (2018)[137] and Apel (2013),[138] the expected utility function,

---

[131] Apel, *supra* note 101, at 69.
[132] *Id*.
[133] *Id*.
[134] Pogarsky, Roche & Pickett, *supra* note 16, at 382.
[135] *Id*.
[136] *Id*. at 381.
[137] *Id*.
[138] Apel, *supra* note 101, at 69.

under which individuals will prefer to violate ICL in response to an international prosecution, may be formalized as follows:

$$U(V) = (1 - \pi_P)U\big(E(G)\big) + \pi_P U\big(E(G) - E(L)\big) > U(A) \quad [1]$$

U(V) is the total expected utility of violating ICL. π_P is the general probability of punishment or sanction for the crime. *U(E(G)–E(L))* is the anticipated (dis)utility of committing an ICL violation, given punishment for such crime. This is a function of the expected gain, *E(G)*, from the crime and the expected loss incurred because of the violation, *E(L)*. In the context of combatants violating ICL, the expected gain may include, but is certainly not limited to, any of the following: monetary gains from pillaging a town; military strength gains from the forced recruitment of child soldiers; territorial, political or economic gains accrued from killing civilians or employing terrorism to promote policy change; deviant psychological pleasure from sexual violence or killing the civilians or individuals *hors de combat* of a perceived enemy outgroup; or reputational gains among fellow combatants accrued by committing these crimes.[139] Individuals are thus assumed to prefer ICL deviance when the expected utility of violating ICL is greater than the expected utility of adherence, *U(A)*, and to prefer adherence when the opposite is true.

To illustrate the function, consider the situation in which a combatant is deciding whether to pillage an undefended city and possibly gain $1,000 in loot. The

[139] *See id.*

combatant judges that the expected probability of getting away with the crime $(1-\pi_p)$ is 50% and thus the expected probability of being punished (π_p) is 50%. The expected punishment includes not only the loss of the $1,000 in loot, but also the loss of all the individual's additional assets, valued at $5,000. The individual's total expected utility from committing the plundering would thus be: *0.5($1,000)+0.5($1,000-$6,000)=-$2,000*. If the utility of adhering to the ICL prohibition of pillage[140] *(U(A))* for the individual is anything greater than -$2,000, then the individual is assumed to prefer to adhere to ICL. In the context of many conflicts where individuals face the risk of execution for not following orders violating ICL, the utility of violation may be far greater than the utility of adherence, despite a negative expected utility of violation.[141]

In accordance with most economic theories of deterrence, the expected loss *(E(L))* of committing an ICL violation may be conceptualized as a function of the certainty, severity and celerity (immediacy) of the penalties anticipated.[142] Given the ambiguity of the effect of the celerity of punishment, many economic models focus on the certainty and severity of sanctions.[143] Between certainty and severity, much more weight is assumed to be given to the certainty of punishment than the severity of a legal sanction, given that the

---

[140] Rome Statute, *supra* note 76, at art. 8(2)(b)(xvi).
[141] *See* HUM. RTS. WATCH, *supra* note 79, at 26.
[142] *See* Daniel S. Nagin, *Deterrence: A Review of the Evidence by a Criminologist for Economists*, 5 ANN. REV. ECON. 83, 85 (2013) [hereinafter Nagin, *A Review of the Evidence*]; Nagin, *Deterrence in the Twenty-First Century*, *supra* note 15, at 205-06.
[143] Nagin, *A Review of the Evidence*, *supra* note 142, at 85.

mere violation of a law, charge, or arrest may all be associated with informal costs, regardless of whether the offender is convicted of the crime and sentenced.[144] Thus, changing the perceived probability of indictment or apprehension is assumed to have more of an effect on individual decision-making regarding ICL adherence than changing the severity of the sanction.[145] The expected loss incurred from committing an ICL violation can therefore be disaggregated as follows:

$$E(L) = E\left(L_V,\, L_I,\, L_A,\, L_C,\, \pi_V,\, \pi_I,\, \pi_A \,\middle|_{I,}\, \pi_C \,|_A \,|_I\right) = \pi_V L_V + \pi_I L_I + \pi_I \pi_{A|_I} L_A + \pi_I \pi_{A|_I} \pi_{C|_{A|_I}} L_C \quad [2]$$

L_V is the informal cost associated with commission of an ICL violation, which may include anything from being wounded in the process, to the psychological costs of guilt or post-traumatic stress disorder (PTSD), or the reputational costs accompanying the naming-and-shaming by a human rights NGO. π_V is the probability of incurring L_V. L_I is the informal cost associated with an indictment, which may include social stigmatization, the loss of material support, and economic sanctions, and π_I is the probability of incurring such loss. L_A is the informal loss associated with apprehension, which may include the loss of economic assets and political standing, and $\pi_{A \mid I}$ is the conditional probability of incurring this loss given indictment. L_C is the formal legal cost associated with conviction, or the loss incurred from a sentence to either a monetary fine or being imprisoned, and $\pi_{C \mid A \mid I}$ is the

[144] *Id*. at 85-86.
[145] *See id*. at 86 (formalizing this argument).

conditional probability of conviction given apprehension and indictment.[146]

The convenience of this model is that it allows for the precise explanation of the mechanisms connecting international prosecutions to individual System 2 judgments.[147] First, prosecutorial deterrence assumes that individuals will generally be deterred from committing ICL violations if they judge the expected formal legal costs associated with such violations (L_C) and the associated probability of incurring such costs $(\pi_I \pi_{A \mid I}\ \pi_{C \mid A \mid I})$ as high relative to the gains to be made by the violations $(E(G))$ and thus, $U(V) < U(A)$. The greater weight given to the certainty of punishment in the model favors a policy focused on increasing the indictment (π_I), apprehension $(\pi_{A \mid I})$, and conviction risk $(\pi_{C \mid A \mid I})$, to increase the value of the expected loss, promoting deterrence.[148]

Specific prosecutorial deterrence should be more effective than general, given that under the model, individuals who are actually subject to international prosecutions, or are indicted, should perceive a greater expected loss in comparison to individuals who have not been indicted, all else being equal.[149] Moreover, international criminal tribunals, including the ICC, tend to focus their prosecutions on those individuals most responsible for ICL violations, including government

---

[146] *See id*.
[147] *See supra* Section II.
[148] *See* Nagin, *A Review of the Evidence*, *supra* note 142, at 85-86.
[149] *See* Akhavan, *supra* note 32, at 746.

political elites, military leaders, and rebel commanders.[150] While foot soldiers may be subject to domestic prosecutions, commanders are more likely to be subject to domestic and/or international prosecutions.[151] Accordingly, individuals with more responsibility for military strategy or for subordinate military personnel should generally perceive a greater probability of punishment (π_P) from an international prosecution than foot soldiers, making deterrence of leaders more effective under the model. There is some suggestive evidence for the effect of ICC actions in deterring specifically commanders in the DRC, where warlords reported fearing arrest for their use of child soldiers in the wake of the ICC Lubanga case, leading to the reduction in the use of child soldiers.[152]

Nonetheless, there are several reasons why international prosecutions may lack a prosecutorial deterrent effect on individuals. First, the dehumanization of an enemy outgroup may reduce the pain empathy associated with killing civilian members of such outgroup, thus reducing the psychological cost of committing the ICL violation (L_V) and the overall expected loss $(E(L))$.[153] The evidence noted by Broache (2016) regarding the Hutu FDLR supports this hypothesis, suggesting that killing Tutsis produces a positive

---

[150] *ICC Situations and Cases*, COALITION FOR THE INT'L CRIM. CT., https://www.coalitionfortheicc.org/explore/icc-situations-and-cases (last visited Oct. 6, 2021).
[151] *See id*.
[152] Parmar, *supra* note 124, at 180.
[153] *See* Gail B. Murrow & Richard Murrow, *A Hypothetical Neurological Association between Dehumanization and Human Rights Abuses*, 2 J. L. & BIOSCI. 336, 343 (2015); *infra* Section V.E.

utility for members of the group.[154] Additionally, at the time of writing the ICC has only opened 30 cases, involving 46 defendants.[155] Coupled with its lack of a global police force and its dismal apprehension and conviction record, the probability of individuals being indicted, apprehended, or convicted by the ICC is objectively quite low.[156] The perception of a low probability of punishment by the ICC, indicating a low expected loss incurred from ICL violations limiting the ICC's deterrent effect, has been noted even in situations in which the ICC has initiated prosecutions, including Sudan[157] and Kenya.[158] International prosecutions may also lack a deterrent effect in situations such as that of the LRA, where the failure to follow an order to commit an ICL violation is likely to result in the execution of an individual, rendering the utility of a violation *(U(V))* much greater than the utility of adherence *(U(A))*.[159] Further, international prosecutions may lack a deterrent effect if individuals view arrest and shelter by the international court as producing maximal utility and thus commit ICL violations in order to be

---

[154] *See* Broache, *supra* note 87, at 31-2.
[155] *Defendants*, INT'L CRIM. CT.: SITUATIONS AND CASES, https://www.icc-cpi.int/Pages/defendants-wip.aspx (last visited Oct. 6, 2021).
[156] *See* Olivia Bueno, *Deterrence in Sudan: The Limits of a Lonely Court*, *in* TWO STEPS FORWARD, ONE STEP BACK: THE DETERRENT EFFECT OF INTERNATIONAL CRIMINAL TRIBUNALS, *supra* note 124, at 222, 240.
[157] *Id*. at 230.
[158] Evelyne Asaala, *The Deterrence Effect of the International Criminal Court: A Kenyan Perspective*, *in* TWO STEPS FORWARD, ONE STEP BACK: THE DETERRENT EFFECT OF INTERNATIONAL CRIMINAL TRIBUNALS, *supra* note 124, at 252, 259.
[159] *See* HUM. RTS. WATCH, *supra* note 79, at 26.

arrested.[160] Evidence supporting this phenomenon has been noted in the case of Dominic Ongwen, who reportedly surrendered to the ICC to avoid death in the field.[161]

The expected utility model also captures social deterrence resulting from the perceived informal, extra-legal costs associated with the violation *(L_V)*, indictment *(L_I)*, and apprehension *(L_A)* for an ICL violation, raising the expected loss relative to gain and leading individuals to view ICL adherence as superior to deviance.[162] There is suggestive evidence for the presence of a social deterrence effect in Sudan, where the referral of the Darfur situation to the ICC in 2005, and the consequent focus of the international community's attention on Sudan, was linked to a decrease in violence against civilians in the country.[163] However, any social deterrent effect was reduced by Bashir's instrumental use of the ICC's actions in the Sudanese case, referring to the efforts as imperialistic to garner support for his government among African countries and reduce the informal costs associated with the ICC actions against him.[164]

---

[160] The domestic analogue of this phenomenon may be found in the example of homeless people committing crimes to seek shelter in jails. *See* Randeep Ramesh, *A Fifth of All Homeless People Have Committed a Crime to Get Off the Streets*, THE GUARDIAN (Dec. 22, 2010, 19:01 EST), https://www.theguardian.com/society/2010/dec/23/homeless-committing-crimes-for-shelter.

[161] Associated Press, *Surrendered LRA Officer Says He Didn't Want to Die in Bush*, DAILYMAIL (Jan. 19, 2015, 18:39 EDT), https://www.dailymail.co.uk/wires/ap/article-2916764/Central-African-Republic-rebels-seek-US-reward-Ongwen.html.

[162] *See* Jo & Simmons, *supra* note 87, at 450; Kim & Sikkink, *supra* note 87, at 943.

[163] Bueno, *supra* note 156, at 233-34.

[164] WEGNER, *supra* note 44, at 116-17.

Finally, the expected utility model accounts for the escalation mechanism.[165] In accordance with the mechanism, individuals faced with a prosecution view the utility of ICL deviance *(U(V))* as greater than that of adherence *(U(A))*, given that escalating their ICL violations is associated with the expected gains in their power or territorial control and avoiding prosecutions.[166] In these situations, the expected gains to be made from ICL violations are perceived as high in relation to the expected losses, which, coupled with a likely low perceived probability of punishment, promotes ICL deviance. Nonetheless, in accordance with Bayesian Learning Theory, the values of individuals' expected losses and gains from crime may vary over time with the information they receive.

B. Bayesian Learning Theory

Empirical evidence demonstrates that individuals' perceptions of the probability of punishment by international prosecutions are not stable across time and change with the receipt of new information.[167] Bayesian learning theory accounts for this phenomenon. Bayesian updating models demonstrate that actors considering violating a law maintain a perceived prior probability of punishment, which is updated with the receipt of new information relevant to punishment risk, termed "signals."[168] Individuals ascribe different amounts of weight to a signal, which is then incorporated into a new

---

[165] *See supra* Section II.D.
[166] *See* Kim & Sikkink, *supra* note 87, at 944; Broache, *supra* note 11, at 393; Jo & Simmons, *supra* note 87, at 445.
[167] Pogarsky, Roche & Pickett, *supra* note 23, at 88-89.
[168] *Id*.

posterior probability of punishment.[169] The signal that individuals consider may be perceived as the ratio of the number of arrests for ICL violations to the number of violations committed in a given time period.[170] If an individual is arrested, the value of the signal should increase, increasing the perceived probability of punishment risk in the future.[171] However, if an individual commits an ICL violation and is not arrested, or avoids punishment, the value of the signal should decrease, decreasing the perceived probability of punishment.[172] In accordance with the "experience effect," new signals from arrests or sanction avoidance should be given less weight in calculating the posterior probability of punishment, as individuals become more certain of the perceived probability of punishment.[173] Offenders may consider both personal punishment experience and the punishment experience of an associate in updating their perceived punishment probabilities, but "a personal punishment experience will likely be weighted more heavily than a friend's vicarious punishment experience."[174]

---

[169] *Id*.
[170] Shamena Anwar & Thomas A. Loughran, *Testing a Bayesian Learning Theory of Deterrence among Serious Juvenile Offenders*, 49 CRIMINOLOGY 667, 670 (2011).
[171] *Id*.
[172] *Id*.
[173] *Id*. at 673.
[174] Pogarsky, Roche & Pickett, *supra* note 23, at 90. Anwar and Loughran (2011) provide a helpful Bayesian updating model, under which the probability of apprehension given indictment, $\pi_{A \mid I}$, of Equation 2 above, may be re-written as the posterior probability of apprehension given indictment, consisting of the weighted average of the new information signal and the prior probability of apprehension given indictment: $p_{i,t} = a_{i,t}q_{i,t} + (1 - a_{i,t})p_{i,t-1}$. $p_{i,t}$ is the perceived posterior probability of apprehension

While a study of the effects of international prosecutions explicitly accounting for Bayesian updating has yet to be conducted, there is suggestive evidence that individuals subject to the actions of the ICC update their perceived probabilities of punishment in accordance with the theory. With regard to the LRA, there is evidence that the "LRA was initially not well informed about how the ICC exactly worked" and there were "reports that the LRA leadership expected international forces to come in and arrest them if the ICC issued warrants."[175] After the ICC issued warrants for the arrests of the LRA leaders in 2005, including Joseph Kony, a former combatant stated that "Kony was afraid of being tried by the ICC because he had the mistaken belief that he would be strangled like Saddam Hussein and that's why he kept pushing for the withdrawal of the arrest warrants."[176] However, "[a] former combatant noted that once it became apparent to Kony that the ICC did not have effective mechanisms to enforce the warrants, his fear of being arrested and prosecuted by the ICC diminished and he continued his brutal attacks against civilian populations."[177]

---

given indictment for individual i in time period t. $q_{i,t}$ is the signal of the apprehension rate, which is given weight $a_{i,t}$, and $p_{i,t-1}$ is the perceived prior probability of apprehension given indictment for the prior time period t-1. The signal $q_{i,t}$ may also be re-written as follows: $q_{i,t} = d_{i,t}(A_{i,t}/C_{i,t}) + (1 - d_{i,t})s_{i,t}$. $A_{i,t}$ is the number of times the individual i has been apprehended in time period t. $C_{i,t}$ is the number of ICL violations that individual i committed in time period t. $d_{i,t}$ is the weight given to the arrest-to-violation ratio, and $s_{i,t}$ is the weighted average of the unobservable factors affecting the individual's perception of the arrest rate, such as vicarious arrests of associates and perceived ICL enforcement trends. Anwar & Loughran, *supra* note 170, at 675.

[175] WEGNER, *supra* note 44, at 230.

[176] Kasande, *supra* note 100, at 208.

[177] *Id*. at 211 (citation omitted).

Likewise, regarding the ICC's actions in Sudan, a Sudanese activist observed that "[i]n the beginning, the regime and Bashir and everyone was afraid. When Bashir and others found out that the ICC does not have police or international forces, then they returned to business as usual."[178]

This evidence supports the conclusion that ICL violators with little prior knowledge of the actual probability of punishment by the ICC perceive a relatively high probability of punishment and may be mistaken about the ICC having an effective enforcement mechanism. The experiment that was conducted for this article also supports this conclusion, demonstrating that, on average, subjects perceived the probability of punishment by an international court for war crimes to be much higher than the actual probability of punishment.[179] However, as ICL violators continue to avoid arrests for their crimes, their arrest-to-violation ratio decreases. Coupled with new information regarding the ICC's lack of an effective enforcement mechanism, the information signal reduces their posterior perceived probability of punishment by the court, degrading any deterrent effect of international prosecutions.[180] The mentioned experiment further supports this conclusion, showing that knowledge of IHL, which may be associated with knowledge of ICL enforcement

[178] Bueno, *supra* note 156 at 230 (citation omitted).
[179] Grenzow, *supra* note 126, at *Mean Sanction Risk Perceptions (All Subjects)* (showing that, on average, subjects perceived the probability of charge by an international prosecution to be 58%, the probability of apprehension given charge by an international prosecution to be 57%, and the probability of conviction given apprehension by an international prosecution to be 63%).
[180] *See* Bueno, *supra* note 156, at 230.

mechanisms, was significantly associated with perceptions of lower and thus more realistic probabilities of punishment by international courts.[181] Nonetheless, like Bayesian learning theory, prospect theory may also modify the conclusions drawn from expected utility theory.

C. Prospect Theory

While prospect theory assumes the deliberative consideration of alternative choices, and thus is situated within System 2 judgment,[182] it modifies the assumptions of expected utility theory. First, unlike expected utility theory, prospect theory accounts for the fact that the evaluation of risk is reference-dependent, and whether individuals are expected to be risk averse or risk seeking depends on the status quo or reference point of the individual.[183] In their famous article,[184] Kahneman and Tversky (1979) replaced the utility function of expected utility theory with a value function, which "is concave in the domain of gains, favoring risk aversion" and "convex in the domain of losses, favoring risk seeking."[185] In other words, individuals are more willing to gamble with the

---

[181] Grenzow, *supra* note 126, Table 4.3 (showing that the *IHLScore*, indicating knowledge of IHL, is negatively associated with the perceived probabilities of charge and apprehension by international courts for ICL violations).
[182] *See* Pogarsky, Roche & Pickett, *supra* note 16, at 389.
[183] Kahneman, *supra* note 20, at 1454-56.
[184] Kahneman & Tversky, *Prospect Theory*, *supra* note 19.
[185] Kahneman, *supra* note 20, at 1456; *see also* Kahneman & Tversky, *Prospect Theory*, *supra* note 19, at 279; Tversky & Kahneman, *Advances in Prospect Theory*, *supra* note 19, at 297; Amos Tversky & Daniel Kahneman, *The Framing of Decisions and the Psychology of Choice*, 211 SCI. 453, 454 (1981).

risk of loss to avoid losses than they are to obtain a gain.[186] The function is also steeper in the domain of losses compared to the domain of gains, capturing the fact that individuals are loss-averse, meaning that "losses loom larger than gains," or losses are ascribed greater absolute value than objectively equivalent gains.[187] Moreover, the flattening of the value function at more positive gains or more negative losses captures the diminishing sensitivity to gains and losses, "or the declining impact of additional gains or losses with movement away from a reference point or status quo."[188] Finally, prospect theory provides for the nonlinear weighting of probabilities, accounting for the fact that changes in the perceived probabilities of events away from 0% or towards 100% are weighted much more heavily than changes in perceived probabilities around 50%.[189]

There are several relevant implications of prospect theory for the effects of international prosecutions on individual decision-making. First, diminishing sensitivity to losses is relevant to the effect of increasing sanction severity. Depending on the reference point of the individual and the type of crime committed, increasing prospective sanctions (L_C) to a certain threshold may have a deterrent effect.[190] However,

---

[186] Kahneman, *supra* note 20, at 1456.
[187] Tversky & Kahneman, *Advances in Prospect Theory*, *supra* note 19, at 298.
[188] Pogarsky, Roche & Pickett, *supra* note 16, at 385 (citation omitted); *see also* Tversky & Kahneman, *Advances in Prospect Theory*, *supra* note 19, at 303.
[189] Tversky & Kahneman, *Advances in Prospect Theory*, *supra* note 19, at 313.
[190] *See supra* Section IV.A.

increasing penalties beyond such threshold will have a diminishing impact on deterring ICL deviance.[191] For instance, individuals may be more deterred by an increase in sanction severity from one to two years imprisonment than they would be from an increase from thirty-one to thirty-two years imprisonment.[192] This observation may offer support for those who argue against the death penalty and for its absence from the Rome Statute of the ICC.[193]

Loss aversion is also relevant to the impact of international prosecutions on individuals. Jervis (2017)[194] and Johnson and Tierney (2018/19)[195] note the presence of loss aversion in the foreign policy context. Some political elites, when faced with the prospect of a loss of power or prestige, become risk-seeking and more willing to engage in a risky war, explaining in part Germany's prosecution of World War I.[196] Facing losses in a war, state actors may also double down on their engagement and escalate hostilities, despite the fact that there is a low probability of victory and a high risk of further loss, explaining in part the U.S. escalation in Vietnam.[197] In the context of law enforcement, loss aversion may explain the phenomenon of resisting arrest, which is a highly risky choice, given its low probability of success and the severe sanctions associated with

---

[191] *See* Pogarsky, Roche & Pickett, *supra* note 23, at 385.
[192] *See id*.
[193] *See* Rome Statute, *supra* note 76, art. 77.
[194] ROBERT JERVIS, HOW STATESMEN THINK: THE PSYCHOLOGY OF INTERNATIONAL POLITICS Ch. 4 (2017).
[195] Dominic D.P. Johnson & Dominic Tierney, *Bad World: The Negativity Bias in International Politics*, 43 INT'L SEC. 96, 109 (2018/19).
[196] JERVIS, *supra* note 194, at 94.
[197] Johnson & Tierney, *supra* note 195, at 110.

failure.[198] As Pogarsky, Roche, and Pickett (2018) explain, "[o]nce a police encounter reaches a point at which the suspects perceive they will be arrested, this realization can shift the offender's reference point, such that the impending arrest is a loss thereby triggering the type of risk-seeking behavior that resisting arrest entails."[199]

These examples help illustrate why prospect theory may provide a superior explanation to expected utility theory for the escalation effect of international prosecutions on individual decision-making.[200] Specifically, when faced with the prospect of a potential loss of power, money, prestige, or freedom from being indicted or arrested by an international court, combatants' reference points may be shifted to the domain of losses. Combatants may therefore become risk-seeking, choosing to engage in the highly risky behavior of escalating hostilities and committing ICL violations, to avoid losses from apprehension and prosecution. The empirical findings discussed above support this hypothesis. For instance, when faced with the calls for his arrest, it appears that Bosco Ntaganda's mindset shifted into the domain of losses, which caused him to engage in highly risky behavior—forming the M23 and increasing hostilities and atrocities in the Eastern DRC—in order to avoid arrest.[201] Similarly, after the LRA failed to dodge the ICC warrants for their leaders' arrests through the Juba Talks, the leaders' mindsets were likely shifted into the domain of losses and they chose to engage in

---

[198] Pogarsky, Roche & Pickett, *supra* note 23, at 386.
[199] *Id.*
[200] *See supra* Section II.D.
[201] *See* Broache, *supra* note 11, at 404-05.

the highly risky behavior of increasing hostilities and attacks on civilians towards the end of 2008, to increase their power and avoid arrest.[202]

Criminologists have also observed that nonlinear weighting of probabilities influences offenders' perceptions of the probability of punishment. Specifically, they have found that "the significant negative association between the perceived probability of arrest and offending grew stronger at the upper end of the risk continuum."[203] Relatedly, research supports the finding that people are averse to ambiguity, and "two offenders may treat identical subjective probabilities of arrest differently (e.g., 40%), depending on how accurately they believe their estimate captures the true probability of arrest."[204] This conclusion is relevant to understanding how police crackdowns deter crime, because "[d]eterrence decay can occur as any initial ambiguity about enforcement during the newly instituted crackdown subsides" while "residual deterrence can persist after the crackdown ends because it takes time for people to learn it has ended and thus have their ambiguity correspondingly reduced."[205] This finding supports a policy of rotating crackdowns to promote enforcement ambiguity, "maximize residual deterrence and minimize deterrence decay."[206]

In the context of individuals responding to international prosecutions, ambiguity aversion may offer an alternative explanation to the Bayesian learning theory for deterrence

---

[202] *See* WEGNER, *supra* note 44, at 230-31.
[203] Pogarsky, Roche & Pickett, *supra* note 23, at 386.
[204] *Id.*
[205] *Id.* at 387.
[206] *Id.*

decay discussed above.[207] Specifically, as the cases of Kony[208] and Bashir[209] suggest, individuals with little knowledge of the actual enforcement capacity of an international court may perceive the probability of punishment to be ambiguous and they may provide greater weight to the ambiguous probability of enforcement, promoting the deterrent impact of the court. However, as individuals gain more information about the probability of enforcement, it may become less ambiguous, degrading the deterrent impact of international prosecutions. This observation favors a policy promoting the ambiguity of the probability of punishment for ICL violations, which may support deterrence despite a low objective probability of sanction. Governments may help promote enforcement ambiguity by rotating crackdowns of ICL violations.[210] Nonetheless, like prospect theory, the heuristics and biases of System 1 may modify the assumptions of expected utility theory. This article now turns to the psychological mechanisms of System 1.

IV. SYSTEM 1 HEURISTICS AND BIASES

Several System 1 heuristics and biases are likely to affect individuals' reactions to international prosecutions and bias their System 2 judgments. Specifically, the representativeness heuristic, anchoring heuristic, availability heuristic, affect heuristic, and the dehumanization bias may all alter individual

---

[207] *See supra* Section IV.B.
[208] *See* Kasande, *supra* note 100, at 208, 211.
[209] *See* Bueno, *supra* note 156, at 230.
[210] *See* Pogarsky, Roche & Pickett, *supra* note 23, at 387.

System 2 judgments. As a result, individual reactions to international prosecutions may not align with the predictions of the System 2 theories discussed in the previous section. Therefore, it is imperative that international prosecutors consider whether and how the targets of their prosecutions may be affected by these heuristics and biases, to determine how they might react to ICL enforcement measures. This section will discuss these heuristics and biases in turn and the relevant policy implications that follow from them.

A. The Representativeness Heuristic

The representativeness heuristic and its attendant biases may skew individual System 2 judgments concerning the probability of punishment by international prosecutions. Instead of relying on population probabilities or base rates, individuals may evaluate "subjective probability by the degree of correspondence between the sample and the population, or between an occurrence and a model,"[211] in accordance with the representativeness heuristic. This heuristic produces three relevant systematic biases of System 2 judgments, including: 1) the base rate fallacy or base rate neglect; 2) the gambler's fallacy; and 3) the conjunction fallacy.[212]

The base rate fallacy refers to individuals' tendency to ignore population base rates altogether.[213] The base rate fallacy

---

[211] Daniel Kahneman & Amos Tversky, *Subjective Probability: A Judgment of Representativeness*, 3 COGNITIVE PSYCH. 430, 451 (1972).
[212] MCDERMOTT, *supra* note 121, at 59-63.
[213] *See* Matthew C. Scheider, *Deterrence and the Base Rate Fallacy: An Examination of Perceived Certainty*, 18 JUST. Q. 63, 66 (2001); MCDERMOTT, *supra* note 121, at 59.

thus violates the assumptions of the Bayesian updating and expected utility models discussed above,[214] which assume reliance on the population base rates of the prior probability of punishment. In a survey experiment examining the effect of exposure to objective population base rates on individuals' perceptions of the certainty of arrest for various crimes, Scheider (2001) finds that "[i]ndividuals relied more heavily on objective information when judging others' certainty than when judging personal certainty," supporting the conclusion that "base rates are used more often when other information is lacking."[215] This finding suggests that, if individuals have little prior information concerning the probability of punishment by an international prosecution, they are more likely to rely on any perceived objective information that is provided to them concerning the certainty of punishment. However, if individuals have prior information which they perceive to be relevant to judging the probability of punishment, they may rely on this information instead of an objective base rate.

These findings suggest that prosecutors should signal that there is a high probability of punishment for ICL violations, emphasizing individual cases and successful prosecutions. This may increase individuals' estimates of the probability of punishment, and thus the expected loss incurred from an ICL violation, especially if they have little prior information concerning the probability of punishment. As a result, prosecutors may promote the deterrence of ICL

[214] *See supra* Section IV.A-B.
[215] Scheider, *supra* note 213, at 77.

violations, even if the actual population base rate of punishment is low.

The gambler's fallacy refers to situations in which individuals "impute interdependencies among chance events."[216] Criminological researchers have "found some offenders, particularly those with little offending experience, tended to reduce their perceived arrest risk after experiencing punishment."[217] The inexperienced offenders perceive "that since a rare event has occurred, that rare event is unlikely to recur because things even out."[218] In the context of international prosecutions, if individuals are arrested for ICL violations and released, they may perceive a lower probability of punishment in the future, given that the rare event of being arrested for an ICL violation has already occurred. This may in turn decrease the expected loss that individuals judge they will incur from ICL violations, diminishing the deterrent effect of international prosecutions. As in the case of the base rate fallacy, this suggests that prosecutors should continue to signal a high probability of punishment for ICL violations to individuals who are arrested and released. This may counteract their subjective discounting of the probability of recurrent punishment for prospective future crimes.

The conjunction fallacy refers to situations in which individuals overestimate the probability of an event occurring. Specifically, individuals find more detailed events to be more likely because of their increased plausibility and imaginability,

---

[216] Pogarsky, Roche & Pickett, *supra* note 23, at 392.
[217] *Id.* (citing Pogarsky & Piquero, *Can Punishment Encourage Offending? Investigating the "Resetting" Effect*, 40 J. RSCH. CRIM. & DELINQ. 95 (2003)).
[218] *Id.*

even though "[o]bjectively, the more detailed a hypothetical future [event] is, the less probable it becomes because its occurrence requires the confluence of more circumstances."[219] In the context of international prosecutions, individuals may perceive a higher probability of punishment for ICL violations when they receive more detailed information concerning ICL enforcement measures. This may include information about the entities involved in apprehending war criminals and other details of apprehensions. By increasing the expected probability of punishment, this detailed information may enhance the deterrent effect of international prosecutions.

B. The Anchoring Heuristic

Like the representativeness heuristic, the anchoring heuristic may also alter the assumptions of expected utility theory, skewing individuals' System 2 judgments of the probability of punishment by an international prosecution. The anchoring heuristic "refers to predictions about frequency or likelihood that are based on initial values (anchors) that are then insufficiently adjusted to reach new estimates."[220] Criminological researchers have shown that, in accordance with the anchoring heuristic, providing individuals information about apprehension risk will dramatically influence their apprehension risk estimates.[221] In an experiment, Pogarsky, Roche and Pickett (2017) asked "participants whether their

---

[219] Pogarsky, Roche & Pickett, *supra* note 23, at 92; *see also* KAHNEMAN, *supra* note 101, at 159; MCDERMOTT, *supra* note 121, at 60.
[220] MCDERMOTT, *supra* note 121, at 67.
[221] Pogarsky, Roche & Pickett, *supra* note 23, at 99-100.

apprehension risk was higher or lower than a specific number, which [they] randomized either to be low ('19 percent') or high ('79 percent')."[222] They then asked subjects to estimate the actual apprehension risk. Confirming the anchoring heuristic, participants who were primed with the lower number reported a perceived apprehension risk that was significantly (seventeen percentage points) lower than those who read the high number.[223] These findings suggest that, if primed with information of a high probability of punishment for ICL violations, individuals will adjust their probability estimates toward the high probability estimate. As a result, providing individuals with information indicating a high probability of punishment may increase their perceived expected losses incurred from ICL violations and promote the deterrent effect of international prosecutions.

C. The Availability Heuristic

The availability heuristic may further bias individuals' System 2 judgments of the probability of punishment for ICL violations. Pursuant to this heuristic, "people estimate an event's probability based on whether they can either 1) quickly recall relevant examples or 2) easily imagine a scenario where the event would occur."[224] In other words, instead of trying to determine the population base rate for the probability of an event, individuals determine the probability of an event based on the availability of information concerning the event in their

---

[222] *Id.*
[223] *Id.* at 100.
[224] *Id.* at 93; *see also* Kahneman & Tversky, *supra* note 211, at 452.

memory or imagination. For example, in accordance with the availability heuristic, individuals may experience "the temporary rise in the subjective probability of an accident after seeing a car overturned by the side of the road" or notice "an increase in the subjective probability that an accident or malfunction will start a thermonuclear war after seeing a movie in which such an occurrence was vividly portrayed."[225]

The availability heuristic may therefore bias individuals' System 2 judgments of the probability of punishment for ICL violations by making individuals perceive a higher risk of punishment if they recently view another individual being apprehended for a crime.[226] In contrast, the heuristic may also cause individuals to perceive a lower risk of punishment if they recently observe themselves or another individual avoiding punishment for an ICL violation.[227] These observations suggest that prosecutors should widely disperse information portraying the apprehension and punishment of individuals for ICL violations, especially to individuals known to be at risk of committing a crime. By making information concerning punishment risk more available, this may increase individuals' subjective estimates of the probability of punishment for ICL violations, supporting the deterrent effect of international prosecutions.

---

[225] Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 5 COGNITIVE PSYCH. 207, 230 (1973).
[226] *See* Pogarsky, Roche & Pickett, *supra* note 23, at 93.
[227] *Id.*

D. The Affect Heuristic

The affect heuristic may also alter individuals' System 2 judgments in reaction to international prosecutions. "Affect" is generally defined as "the sense (not necessarily conscious) that something is good or bad."[228] In accordance with the affect heuristic, if an individual's perception of a choice promotes positive affect, or the individual has favorable feelings regarding the choice, the individual is biased to infer the risks of the choice to be low and the benefits to be high.[229] Conversely, if the perception of the choice promotes negative affect, or negative feelings concerning the choice, then the risks of the choice are inferred to be high and the benefits inferred to be low.[230] Researchers have confirmed the presence of the affect heuristic in subjective estimations of sanction risk.[231] These findings suggest that if individuals are primed with negative information concerning ICL violations, they may increase their estimates of the probability of punishment for such violations, promoting the deterrent effect of international prosecutions. A negative prime may include information about

---

[228] Slovic, *supra* note 102, at 82.
[229] Paul Slovic et al., *Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality*, 24 RISK ANALYSIS 311, 315 (2004).
[230] *Id.*
[231] Pogarsky, Roche & Pickett, *supra* note 23, at 101.

the negative consequences that ICL violations have for the victims of such crimes.[232]

Related to the affect heuristic, emotional deterrence may also alter System 2 judgments.[233] In an experiment Pickett, Roche, and Pogarsky (2018) found that fear "mediates the effects of perceived apprehension risk on criminal decision-making" and that "for most offenses evaluated, fear is the strongest predictor of situational intentions to offend."[234] Specifically, they found that perceived fear of apprehension was positively associated with perceived risk of apprehension.[235] In turn, perceived fear of apprehension mediated the negative relationship between perceived apprehension risk and intentions to offend.[236] Similarly, in the context of international prosecutions, perceived fear associated with ICL violations may mediate the negative relationship between the perceived probability of punishment and individuals' willingness to commit crimes. Perceived fear of ICL violations, moreover, may itself be negatively associated with individual intentions to violate ICL.

These conclusions have incredibly important policy implications for the effect of international prosecutions in deterring ICL violations. Given resource constraints and the lack of an effective enforcement mechanism, it is difficult for international tribunals such as the ICC to promote an

---

[232] *Id.*
[233] *See* Justin T. Pickett, Sean Patrick Roche & Greg Pogarsky, *Toward a Bifurcated Theory of Emotional Deterrence*, 56 CRIMINOLOGY 27 (2018); *see also* Section IV.
[234] Pickett, Roche & Pogarsky, *supra* note 233, at 45.
[235] *Id.* at 39-40.
[236] *Id.* at 42-3.

objectively high probability of punishment for crimes under their jurisdiction. However, the findings concerning emotional deterrence suggest that "any policy that increases fear of apprehension may be a viable deterrence strategy, regardless of whether it impacts the objective level of sanction risk."[237] By priming fear of apprehension or fear associated with ICL violations, prosecutors may promote emotional deterrence without necessarily increasing ICL enforcement. Thus, aside from communicating a high probability of punishment, prosecutors should transmit signals to individuals which associate fear with contemplated ICL violations. Such signals may include vivid and emotional stories concerning the negative consequences of crimes for victims and perpetrators.[238]

E. The Dehumanization Bias

The infrahumanization and dehumanization biases, commonly involved in conflict contexts,[239] may also alter individuals' System 2 judgments, reducing the deterrent effect of international prosecutions and inhibiting them from socializing individuals towards ICL adherence. "Infrahumanization" is the implicit cognitive bias through which humans "are inclined to perceive members of outgroups as somewhat less human, or more animal-like, than

---

[237] *Id.* at 45.
[238] *Id.*
[239] *See* Daniel Bar-Tal, *Causes and Consequences of Delegitimization: Models of Conflict and Ethnocentrism*, 46 J. SOC. ISSUES 65 (1990).

themselves."[240] "Dehumanization" is the magnification of this bias to the point where outgroup members are no longer perceived as human.[241]

It is hypothesized that the dehumanization bias facilitates the moral exclusion of outgroups, involving the perception of an outgroup as undeserving of the application of moral values and norms, including ICL and IHL norms.[242] Bar-Tal (1990) hypothesized that dehumanization and moral exclusion are deployed to justify atrocities of an ingroup against an outgroup.[243] Specifically, the perception of an outgroup threat causes the ingroup to attack and commit atrocities against the outgroup, which then commits atrocities against the ingroup.[244] This causes the ingroup to dehumanize the outgroup in order to justify and explain ICL violations on both sides.[245] Murrow and Murrow (2015) similarly hypothesize that dehumanization occurs as a defense mechanism to reduce empathy toward an outgroup, in order to facilitate the commission of violence against members of that group.[246] As a result of

---

[240] Jacques-Philippe Leyens et al., *Infra-humanization: The Wall of Group Differences*, 1 SOC. ISSUES & POL'Y REV. 139, 143 (2007).
[241] *Id.*
[242] *See* Susan Opotow, *Moral Exclusion and Injustice: An Introduction*, 46 J. SOC. ISSUES 1 (1990); Bar-Tal, *supra* note 239, at 65; Lasana T. Harris & Susan T. Fiske, *Dehumanized Perception: A Psychological Means to Facilitate Atrocities, Torture, and Genocide?*, 219 ZEITSCHRIFT FÜR PSYCHOLOGIE [JOURNAL OF PSYCHOLOGY] 175, 180 (2011).
[243] Bar-Tal, *supra* note 239, at 72-73.
[244] *Id.*
[245] *Id.*
[246] Murrow & Murrow, *supra* note 153, at 343.

dehumanization, ICL violations become less emotionally upsetting.[247]

Several psychological experiments have supported these hypotheses. For instance, studies have found that the infrahumanization bias is primed by exposure to outgroup photographs and names, along with pictures associated with human violent acts (e.g., kidnappings, weapons, and street gangs).[248] Castano and Giner-Sorolla (2006) also found that, when subjects were made aware of their ingroup's mass killing of an outgroup, they infrahumanized the victims of such killing more.[249] Viki, Osgood and Phillips (2013) established a link between the dehumanization of an outgroup and subjects' willingness to torture outgroup members.[250] Maoz and McCauley (2008) similarly showed a connection between the perception of an outgroup threat, dehumanization, and support for policies violating the human rights of an outgroup.[251]

The dehumanization bias may thus skew individuals' System 2 judgments in reaction to international prosecutions in at least three important ways. First, by promoting the moral

---

[247] *See* G. Tendayi Viki, Daniel Osgood & Sabine Phillips, *Dehumanization and Self-Reported Proclivity to Torture Prisoners of War*, 49 J. EXPERIMENTAL SOC. PSYCH. 325, 326 (2013).

[248] *See* Leyens et al., *supra* note 240, at 147; Naira Delgado et al., *Priming Effects of Violence on Infrahumanization*, 12 GRP. PROCESSES & INTERGRP. REL. 699, 703 (2009); Naira Delgado Rodríguez et al., *Contextual Variations of Infrahumanization: The Role of Physical Context and Territoriality*, 34 BASIC & APPLIED SOC. PSYCH. 456, 459-60 (2012).

[249] Emanuele Castano & Roger Giner-Sorolla, *Not Quite Human: Infrahumanization in Response to Collective Responsibility for Intergroup Killing*, 90 J. PERSONALITY & SOC. PSYCH. 804, 804 (2006).

[250] Viki, Osgood & Phillips, *supra* note 247, at 325-28.

[251] Ifat Maoz & Clark McCauley, *Threat, Dehumanization, and Support for Retaliatory Aggressive Policies in Asymmetric Conflict*, 52 J. CONFLICT RESOL. 93, 111 (2008).

exclusion of enemy outgroups, the dehumanization bias may cause individuals to perceive that the provisions of IHL and ICL do not apply to their interactions with the outgroup. With reference to Equation 2 above,[252] this may reduce the value of their expected loss *(E(L))* incurred from committing an ICL violation, by eliminating the perception of the losses associated with indictment $(\pi_I L_I)$, apprehension $(\pi_I \pi_{A \mid I} L_A)$, and conviction $(\pi_I \pi_{A \mid I} \pi_{C \mid A \mid I} L_C)$ for ICL violations by an international prosecution. In other words, dehumanization may make the individuals think that they are not violating ICL with regards to an enemy outgroup, when they in fact are, and may lead them to not foresee that they will incur costs from a prospective international prosecution. Second, dehumanization may reduce the pain empathy associated with the commission of an ICL violation, diminishing the expected psychological costs of guilt and thus the informal cost (L_V) associated with the crime. The overall reduction of the expected loss incurred from the ICL violation *(E(L))* due to dehumanization may thus increase the expected utility of the violation *(U(V))* in relation to the utility of ICL adherence *(U(A))*. As a result, international prosecutions may not have a deterrent effect in conflicts involving dehumanization. Third, dehumanization may prevent international prosecutions from socializing individuals towards ICL adherence, by causing individuals to fail to perceive ICL and IHL as applicable to their interactions with the enemy outgroup despite knowledge of prosecutions.[253]

---

[252] *See supra* Section IV.A.
[253] *See* Broache, *supra* note 7, at 22.

Broache (2016) notes evidence supporting this conclusion with respect to the Hutu FDLR in the DRC.[254]

Nonetheless, there may be ways to reduce the dehumanization bias and promote the deterrence and socialization effects of international prosecutions. In a functional magnetic resonance imaging (fMRI) study, Harris and Fiske (2007) found that, when subjects were primed to infer the preferences of a dehumanized target, requiring them to mentalize or consider the contents of the target actors' minds through a social individuation task,[255] the subjects' medial prefrontal cortex (mPFC), associated with the perception of humanized target individuals, became more active.[256] This finding demonstrates that mentalization may mitigate the dehumanization bias.[257] Mentalization may mitigate dehumanization and associated moral exclusion by increasing pain empathy for the outgroup[258] and/or by promoting the perception of similarity and connectedness between the ingroup and the outgroup.[259] While it may be difficult to implement in the conflict contexts that give rise to international prosecutions, these findings support a policy of promoting the mentalization by individuals of the enemy outgroup or the perception of similarity and connectedness between the

---

[254] *Id.* at 31-32.
[255] *See* Mary E. Wheeler & Susan T. Fiske, *Controlling Racial Prejudice: Social-Cognitive Goals Affect Amygdala and Stereotype Activation*, 16 PSYCH. SCI. 56, 58 (2005) (discussing the social individuation task); Lasana T. Harris & Susan T. Fiske, *Social Groups that Elicit Disgust Are Differentially Processed in mPFC*, 2 SOC. COGNITIVE & AFFECTIVE NEUROSCIENCE 45, 46 (2007) (discussing mentalization).
[256] Harris & Fiske, *supra* note 255, at 48.
[257] *Id.*
[258] *See* Murrow & Murrow, *supra* note 153, at 343.
[259] *See* Opotow, *supra* note 242, at 7; Leyens et al., *supra* note 240, at 164.

ingroups and outgroups, to reduce the dehumanization bias. By adopting policies that aim to diminish dehumanization, prosecutors may promote the deterrent and socialization effects of international prosecutions.

CONCLUSION AND POLICY IMPLICATIONS

To provide a more comprehensive and realistic theory of individual decision-making in response to international prosecutions, this article drew on advances in criminology, psychology, and behavioral economics, explaining individual decision-making in terms of the two systems of cognition derived from dual-process decision-making theory. In so doing, the article provides a more refined explanation of the mechanisms connecting international prosecutions to individual conduct. At a general level, individual decision-making in response to international prosecutions is mediated by the rationalist Logic of Expected Consequences (LEC) and the norms-based Logic of Appropriateness (LOA). While the expected utility theories underlying the deterrence and escalation mechanisms of international prosecutions assume individual decision-making through the LEC, the socialization mechanism functions through the LOA. However, the logics are not mutually exclusive, and in many situations, they may interact in mediating individual responses to international prosecutions. In turn, System 1—the unconscious, intuitive, emotional, and habit-based cognitive system—explains individual judgment under the unconscious LOA, and thus explains the socialization mechanism connecting international prosecutions to individual decision-making. Meanwhile,

System 2—the conscious and deliberative system—explains conscious judgment under both the LOA and the LEC. Depending on the circumstances facing individuals, System 1 and System 2 may interact in different ways, with System 2 modifying System 1 unconscious judgments in situations where individuals have more time to make a decision or merely ratifying System 1 judgments in situations involving limited time.

Beyond a psychological explanation of the LOA and LEC, applying the insights of behavioral economics to the context of international criminal justice provides a more thorough theoretical development of individual decision-making in response to international prosecutions. The article formalized individual System 2 judgments in response to international prosecutions, demonstrating how the probability of indictment and apprehension are weighed more heavily in expected utility judgments than the severity of sanction costs. The article also discussed the implications of Bayesian updating models and how prospect theory may modify the assumptions of expected utility theory. The article further explained how the representativeness heuristic may bias individuals' rational judgments, in accordance with the base rate, gambler's, and conjunction fallacies. Likewise, it discussed how the anchoring, availability, and affect heuristics, and the dehumanization bias may also skew individuals' System 2 judgments in response to international prosecutions.

Several policy implications may be derived from the findings discussed in this article. In general, the article demonstrates that, in response to international prosecutions, certain environmental or psychological factors may cause individuals to decrease or escalate their ICL violations. To

determine what effects ICL enforcement measures—including indictments, arrests, and convictions—will have on individual conduct, it is imperative that prosecutors consider the psychological circumstances and potential judgments of the individuals involved in a situation under review. This may give prosecutors the ability to better predict the potential effects of their actions and promote the purpose of ICL to prevent crime. This article offers a preliminary framework to use when considering how individuals may react to prosecutions.

Many of the findings discussed also support the conclusion that prosecutors and policymakers should increase their efforts to educate individuals about the illegality of ICL violations. Considering the availability heuristic, individuals who are informed about ICL, especially in an emotionally involving way, may be more likely to remember ICL provisions and comply with them in the future. However, the survey evidence presented in this article poses a policy paradox, because it shows that increased knowledge of international law is also associated with a decrease in the perceived probability of punishment by international courts for war crimes. This finding is in accordance with the Bayesian updating models and the theory of ambiguity aversion, providing that when individuals receive more information regarding the objective probability of punishment for war crimes, they may update their perceptions to a lower probability of punishment, in line with reality, and become less averse to the information as it becomes less ambiguous. In turn, this leads to the decay of the initial deterrence effects of international prosecutions, resulting from the adjustment from higher and more ambiguous perceived probabilities of punishment for ICL violations. The most

obvious way to combat deterrence decay is to increase the objective probability of punishment for ICL violations by increasing the number of charges, arrests, and convictions by international courts and tribunals. The expected utility model developed in this article supports the conclusion that individuals will weigh the probability of punishment the most in their calculations and changes in sanction risk will thus have the most effect in deterring them from violating ICL. Yet, given the resource constraints of international courts, this policy is likely untenable.

A more feasible path to promote the deterrence of ICL violations is for states and international institutions to support domestic enforcement measures, by domestic war crimes courts or military court-martials. More domestic enforcement may lead to an increase in the perceived probability of punishment for individuals, promoting System 2 deterrence. Rotating domestic crackdowns on ICL violators may also promote enforcement ambiguity, which will instigate more aversion to crime and deterrence.

Even in the absence of an objectively high probability of punishment, prosecutors may promote deterrence by signaling to would-be ICL violators a high perceived risk of punishment. The anchoring heuristic supports the conclusion that, even if this information is far from accurate, it may still bias individuals' risk perceptions upwards toward the primed probability estimate. Moreover, the conjunction fallacy supports the provision of detailed information concerning punishment risk, which will promote higher risk estimates, and the emotional deterrence literature supports a policy of associating fear with contemplated ICL violations. In accordance with expected utility and prospect theory,

increasing sanction severity for ICL violations may also promote deterrence, but diminishing sensitivity to losses may make increasing sanctions past a certain threshold ineffective, weighing against the use of the death penalty. Considering the findings concerning the dehumanization bias, prosecutors and other relevant actors should promote the mentalization by would-be ICL violators of the enemy outgroup, or the perception of similarity and connectedness between the ingroups and outgroups involved in a conflict. This may reduce the dehumanization bias and as a result promote the deterrence and socialization mechanisms of international prosecutions.

Nonetheless, more empirical research is needed to confirm the impact of the various heuristics and biases discussed in this article, in the context of individual reactions to international prosecutions. Future research should further explore the mechanisms of social deterrence and individual perceptions of the formal legal costs and the informal extralegal costs associated with prospective ICL violations. Hopefully this article offers another step towards promoting policies that will prevent the commission of atrocities in the future and thus fulfill the purpose of international prosecutions.