

REVISTA BRASILEIRA DE POLÍTICAS PÚBLICAS
BRAZILIAN JOURNAL OF PUBLIC POLICY

Cérebros que punem: uma
revisão crítica da neurociência da
punição

Punitive brains: a critical
review of the neuroscience of
punishment

Ricardo de Lins e Horta

VOLUME 8 • Nº 2 • AGO • 2018
DOSSIÊ ESPECIAL: INDUÇÃO DE COMPORTAMENTOS
(NEUROLAW): DIREITO, PSICOLOGIA E NEUROCIÊNCIA

Sumário

EDITORIAL	24
O Direito na fronteira da razão: Psicologia, neurociência e economia comportamental	24
Patrícia Perrone Campos Mello e Sergio Nojiri	
I. NEURODIREITO: COGNIÇÃO, EMOÇÃO, JUÍZOS MORAIS E CIÊNCIA	26
PENSAR DIREITO E EMOÇÃO: UMA CARTOGRAFIA	28
Nevita Maria Pessoa de Aquino Franca Luna	
NEURODIREITO: O INÍCIO, O FIM E O MEIO	49
Carlos Marden e Leonardo Martins Wykrota	
ENSAIO JURÍDICO SOBRE A RACIONALIDADE HUMANA: MAIORES, CAPAZES E IRRACIONAIS	65
André Perin Schmidt Neto e Eugênio Facchini Neto	
DIVERGÊNCIAS DE PRINCÍPIO: ARGUMENTOS JURÍDICOS E MORAIS EM UM CENÁRIO DE DESACORDOS SOCIAIS	90
André Matos de Almeida Oliveira, Pâmela de Rezende Côrtes e Leonardo Martins Wykrota	
CONSILIÊNCIA E A POSSIBILIDADE DO NEURODIREITO: DA DESCONFIANÇA À RECONCILIAÇÃO DISCIPLINAR	117
Thaís de Bessa Gontijo de Oliveira e Renato César Cardoso	
MODELOS DE MORALIDADE	144
Molly J. Crockett	
A INFELIZ BUSCA POR FELICIDADE NO DIREITO	154
Úrsula Simões da Costa Cunha Vasconcellost, Noel Struchiner e Ivar Hannikainen	
ALÉM DA LIBERDADE: PERSPECTIVAS EM NIETZSCHE	178
Lucas Costa de Oliveira	
A MEDIAÇÃO DE CONFLITOS SOB A PERSPECTIVA DO DESENVOLVIMENTO HUMANO: AS CONTRIBUIÇÕES DA PSICOLOGIA POSITIVA	193
Simone de Biazzi Ávila Batista da Silveira e Deise Brião Ferraz	
NEUROIMAGIOLOGIA E AVALIAÇÃO DE RESPONSABILIDADE	213
Nicole A. Vincent	

II. NUDGES: INDUÇÃO DE COMPORTAMENTOS E POLÍTICAS PÚBLICAS	233
ANÁLISE CRÍTICA DA ORIENTAÇÃO DE CIDADÃOS COMO MÉTODO PARA OTIMIZAR DECISÕES PÚBLICAS POR MEIO DA TÉCNICA NUDGE.....	235
Luciana Cristina Souza, Karen Tobias França Ramos e Sônia Carolina Romão Viana Perdigão	
POLÍTICAS PÚBLICAS E O DEVER DE MONITORAMENTO: “LEVANDO OS DIREITOS A SÉRIO”	252
Ana Paula de Barcellos	
NUDGES E POLÍTICAS PÚBLICAS: UM MECANISMO DE COMBATE AO TRABALHO EM CONDIÇÃO ANÁLOGA À DE ESCRAVO	267
Amanda Carolina Souza Silva, Débhora Renata Nunes Rodrigues e Saul Duarte Tibaldi	
REDUZINDO A TRIBUTAÇÃO COGNITIVA: LIÇÕES COMPORTAMENTAIS PARA A DIMINUIÇÃO DOS EFEITOS PSICOLÓGICOS ADVERSOS DA POBREZA.....	288
Leandro Novais e Silva, Luiz Felipe Drummond Teixeira, Gabriel Salgueiro Soares e Otávio Augusto Andrade Santos	
POLÍTICAS PÚBLICAS EM SUICÍDIO: DO PATERNALISMO CLÁSSICO AO PATERNALISMO LIBERTÁRIO E NUDGING	327
Davi de Paiva Costa Tangerino, Gabriel Cabral e Henrique Olive	
NUDGES COMO POLÍTICA PÚBLICA PARA AUMENTAR O ESCASSO NÚMERO DE DOADORES DE ÓRGÃOS PARA TRANSPLANTE	369
Roberta Marina Cioatto e Adriana de Alencar Gomes Pinheiro	
OS PROGRAMAS DE INTEGRIDADE PARA CONTRATAÇÃO COM A ADMINISTRAÇÃO PÚBLICA ESTADUAL: NUDGE OU OBRIGAÇÃO LEGAL? UM OLHAR SOBRE AS DUAS PERSPECTIVAS	386
Cíntia Muniz Rebouças de Alencar Araripe e Raquel Cavalcanti Ramos Machado	
PATERNALISMO LIBERTÁRIO E PROTEÇÃO JURÍDICA DO AMBIENTE: POR QUE PROTEGER O AMBIENTE TAMBÉM DEVE SER PROTEGER AS LIBERDADES?	406
Mariana Carvalho Victor Coelho e Patryck de Araujo Ayala	
POLÍTICAS PÚBLICAS BASEADAS EM EVIDÊNCIAS COMPORTAMENTAIS: REFLEXÕES A PARTIR DO PROJETO DE LEI 488/2017 DO SENADO	429
Pâmela de Rezende Côrtes, André Matos de Almeida Oliveira e Fabiano Teodoro de Rezende Lara	
III. ECONOMIA COMPORTAMENTAL: VIESES COGNITIVOS E POLÍTICAS PÚBLICAS	455
ECONOMIA COMPORTAMENTAL E DIREITO: A RACIONALIDADE EM MUDANÇA	457
Marcia Carla Pereira Ribeiro e Victor Hugo Domingues	
VIESES COGNITIVOS E DESENHO DE POLÍTICAS PÚBLICAS.....	473
Benjamin Miranda Tabak e Pedro Henrique Rincon Amaral	

A NEUROCIÊNCIA DA MORALIDADE NA TOMADA DE DECISÕES JURÍDICAS COMPLEXAS E NO DESENHO DE POLÍTICAS PÚBLICAS	493
Erik Navarro Wolkart	
DESVIO DE CARÁTER OU SIMPLEMENTE HUMANO? ECONOMIA COMPORTAMENTAL APLICADA AO COMPORTAMENTO DESONESTO.....	524
Diana Orghian, Gabriel Cabral, André Pinto e Alessandra Fontana	
POLÍTICAS PÚBLICAS E A CONCRETIZAÇÃO DE DIREITOS SOCIAIS: TOMADA DE DECISÃO, ARQUITETURA DE ESCOLHAS E EFETIVIDADE	543
Ana Elizabeth Neirão Reymão e Ricardo dos Santos Caçapietra	
BEHAVIORAL ECONOMICS E DIREITO DO CONSUMIDOR: NOVAS PERSPECTIVAS PARA O ENFRENTAMENTO DO SUPERENDIVIDAMENTO	568
Samir Alves Daura	
A EDUCAÇÃO FORMAL PARA O CONSUMO É GARANTIA PARA UMA PRESENÇA REFLETIDA DO CONSUMIDOR NO MERCADO? UMA ANÁLISE COM BASE NA BEHAVIORAL LAW AND ECONOMICS (ECONOMIA COMPORTAMENTAL)	600
Marcia Carla Pereira Ribeiro e Edson Mitsuo Tiujo	
LIBET, DETERMINISMO E CONSUMO: AS INFLUÊNCIAS DO MARKETING E A RELEVÂNCIA DA DELIBERAÇÃO CONSCIENTE NA SUPERAÇÃO CONDICIONAL DE HÁBITOS DE CONSUMO PERIGOSOS	616
Émilien Vilas Boas Reis e Leonardo Cordeiro de Gusmão	
CIÊNCIA DO DIREITO TRIBUTÁRIO, ECONOMIA COMPORTAMENTAL E EXTRAFISCALIDADE.....	640
Hugo de Brito Machado Segundo	
IV. COMPORTAMENTO JUDICIAL: INFLUÊNCIA DE FATORES EXTRAJURÍDICOS	660
FATORES METAPROCESSUAIS E SUAS INFLUÊNCIAS PARA A FORMAÇÃO DA DECISÃO JUDICIAL	662
Rogério Roberto Gonçalves de Abreu, Lúcio Grassi de Gouveia e Virgínia Colares	
“A VIDA COMO ELA É”: COMPORTAMENTO ESTRATÉGICO NAS CORTES.....	689
Patrícia Perrone Campos Mello	
A COMPOSIÇÃO DO ÓRGÃO COLEGIADO E SEUS EFEITOS NA TOMADA DE DECISÃO	720
André Garcia Leão Reis Valadares	
DAS 11 ILHAS AO CENTRO DO ARQUIPÉLAGO: OS SUPERPODERES DO PRESIDENTE DO STF DURANTE O RECESSO JUDICIAL E FÉRIAS	741
José Mário Wanderley Gomes Neto e Flávia Danielle Santiago Lima	

RAZÃO, EMOÇÃO E DELIBERAÇÃO: AS ADEQUAÇÕES REGIMENTAIS DO SUPERIOR TRIBUNAL DE JUSTIÇA PARA A FORMAÇÃO DE PRECEDENTES EFICAZES	758
Peter Panutto e Lana Olivi Chaim	
HEURÍSTICA DE ANCORAGEM E FIXAÇÃO DE DANOS MORAIS EM JUIZADOS ESPECIAIS CÍVEIS NO RIO DE JANEIRO: UMA NOVA ANÁLISE	778
Fernando Leal e Leandro Molhano Ribeiro	
LA PROTECCIÓN DE LOS DERECHOS POLÍTICOS FRENTE A LAS FUNCIONES DISCIPLINARIAS DE LAS AUTORIDADES ADMINISTRATIVAS: SUBSIDIARIEDAD Y DEFERENCIA EN EL SISTEMA INTERAMERICANO DE DERECHOS HUMANOS.....	801
Jorge Ernesto Roa Roa	
V. A INFLUÊNCIA DO GÊNERO NO PROCESSO DECISÓRIO JUDICIAL.....	824
COMO OS JUÍZES DECIDEM OS CASOS DE ESTUPRO? ANÁLISANDO SENTENÇAS SOB A PERSPECTIVA DE VIESES E ESTEREÓTIPOS DE GÊNERO	826
Gabriela Perissinotto de Almeida e Sérgio Nojiri	
GÊNERO E COMPORTAMENTO JUDICIAL NO SUPREMO TRIBUNAL FEDERAL: OS MINISTROS CONFIAM MENOS EM RELATORAS MULHERES?.....	855
Juliana Cesario Alvim Gomes, Rafaela Nogueira e Diego Werneck Arguelhes	
HÉRCULES, HERMES E A PEQUENA SEREIA: UMA REFLEXÃO SOBRE ESTEREÓTIPOS DE GÊNERO, SUBPRESENTAÇÃO DAS MULHERES NOS TRIBUNAIS E (I)LEGITIMIDADE DEMOCRÁTICA DO PODER JUDICIÁRIO.....	878
Jane Reis Gonçalves Pereira e Renan Medeiros de Oliveira	
PRISÃO CAUTELAR DE GESTANTES: ANÁLISE DO FUNDAMENTO FILOSÓFICO DA DECISÃO DO HABEAS CORPUS N. 143.641	912
Artur César Souza e Giovania Tatibana de Souza	
VI. NEURODIREITO APLICADO AO DIREITO E AO PROCESSO PENAL.....	926
CÉREBROS QUE PUNEM: UMA REVISÃO CRÍTICA DA NEUROCIÊNCIA DA PUNIÇÃO	928
Ricardo de Lins e Horta	
A INTUIÇÃO DO DOLO EM DIREITO PENAL: CORRELATOS NEURAIS DA TEORIA DA MENTE, RACIOCÍNIO INDUTIVO E A GARANTIA DA CONVICÇÃO JUSTIFICADA.....	946
Thiago Dias de Matos Diniz e Renato César Cardoso	
AS COMUNIDADES EPISTÊMICAS PENAIS E A PRODUÇÃO LEGISLATIVA EM MATÉRIA CRIMINAL.....	961
Stéphane Enguéléguélé	

DELINQUÊNCIA JUVENIL: RELAÇÕES ENTRE DESENVOLVIMENTO, FUNÇÕES EXECUTIVAS E COMPORTAMENTO SOCIAL NA ADOLESCÊNCIA	980
André Vilela Komatsu, Rafaelle CS Costa e Marina Rezende Bazon	
LÍMITES TEMPORALES A LAS PENAS PRIVATIVAS DE LIBERTAD ATENDIENDO AL DESARROLLO PSICOSOCIAL.....	1001
Silvio Cuneo Nash	
NEUROLAW E AS PERSPECTIVAS PARA UMA ANÁLISE OBJETIVA DO COMPORTAMENTO SUGESTIONADO: REPERCUSSÃO DAS FALSAS MEMÓRIAS NA ESFERA PENAL.....	1017
Mariana Dionísio de Andrade, Marina Andrade Cartaxo e Rafael Gonçalves Mota	
A FALIBILIDADE DA MEMÓRIA NOS RELATOS TESTEMUNHAIS AS IMPLICAÇÕES DAS FALSAS MEMÓRIAS NO CONTEXTO DOS CRIMES CONTRA A DIGNIDADE SEXUAL	1036
Caroline Navas Viana	
A (IR)REPETIBILIDADE DA PROVA PENAL DEPENDENTE DA MEMÓRIA: UMA DISCUSSÃO COM BASE NA PSICOLOGIA DO TESTEMUNHO.....	1058
William Weber Ceconello, Gustavo Noronha de Avila e Lilian Milnitsky Stein	

Cérebros que punem: uma revisão crítica da neurociência da punição*

Punitive brains: a critical review of the neuroscience of punishment

Ricardo de Lins e Horta** ***

RESUMO

Os profissionais envolvidos com o sistema penal, rotineiramente, tomam decisões sobre a culpabilidade e a punibilidade de transgressores. O comportamento de punir um ofensor tem sido estudado por psicólogos e neurocientistas, que buscam compreender quais as motivações subjacentes ao ato de punir, bem como quais as estruturas cerebrais envolvidas no processamento de informações sobre a intencionalidade e gravidade de uma ofensa. Estudos recentes apontam que a decisão de punir envolve áreas cerebrais específicas, cuja origem estaria ligada à funcionalidade, no contexto de evolução da espécie, de estabilizar a cooperação em grupos, por intermédio do reforço das normas sociais mediante a aplicação de sanções. Neste artigo, buscamos fazer uma revisão crítica dessa literatura, apontando que, apesar da importância dos estudos da chamada Neurociência da Punição, os pressupostos por eles adotados ainda são discutíveis, do ponto de vista da biologia evolucionista e da neurociência cognitiva.

Palavras-chave: Acesso à justiça. Neurodireito. Psicologia da punição. Teorias da pena. Retributivismo.

ABSTRACT

Professionals involved in the criminal justice system routinely make decisions about the culpability of offenders. The behavior of punishing an offender has been studied by psychologists and neuroscientists, who seek to understand the motivations underlying the act of punishing, as well as which brain structures are involved in processing information about the intentionality and severity of an offense. Recent studies indicate that the decision to punish involves specific brain areas related to the function, in the context of evolution of the species, to stabilize cooperation in groups, through the reinforcement of social norms due to the application of sanctions. In this article, we seek to make a critical review of this literature, pointing out that, despite the importance of the studies of the so-called Neuroscience of Punishment, its assumptions are still debatable from the point of view both of Evolutionary Biology and Cognitive Neuroscience.

Keywords: *Neurolaw*. Psychology of Punishment. Theories of punishment. Retributivism.

* Recebido em 31/05/2018
Aprovado em 02/07/2018

** Doutorando em Direito (UnB), Mestre em Neurociências e Graduado em Direito (UFMG). Foi pesquisador visitante na École Normale Supérieure, em Paris (2016-2017). Integrante da carreira de Especialista em Políticas Públicas e Gestão Governamental (EPPGG) federal, é atualmente Coordenador-Geral de Articulação Institucional da Escola Nacional de Administração Pública (ENAP).
Email: ricardolinshorta@gmail.com

*** Versão anterior deste trabalho foi apresentada como conferência no III Seminário de Direito, Psicologia e Neurociência, organizado pelo grupo de estudos em Direito, Psicologia e Neurociência (DIPSIN) da Faculdade de Direito de Ribeirão Preto da USP, em 17/10/2017. Agradecemos os valiosos comentários do professor Noel Struchiner (PUC/RJ), que contribuíram para o esclarecimento de diversos pontos do manuscrito original.

1. INTRODUÇÃO

O que se passa na mente de uma testemunha de um crime, ou de alguém que presencia o sofrimento de uma vítima de um delito? Por que e como alguém decide que um violador de normas sociais deve ser punido? Esses são problemas sobre os quais cientistas vêm se debruçando recentemente, utilizando avançadas técnicas experimentais.

Nas últimas décadas, as Ciências Comportamentais vêm adotando um conjunto de métodos de pesquisa que permitem traçar um cenário cada vez mais preciso desse complexo fenômeno que é o comportamento humano. Explicações biopsicossociais hoje são corriqueiras tanto para os comportamentos comumente identificados com os demais animais não humanos, como a agressão e o desejo sexual; quanto para aqueles que um dia se sustentou, equivocadamente, serem exclusivos da espécie humana, como a linguagem, as relações de poder ou os comportamentos altruístas e cooperativos¹. Nesse sentido, os resultados mais promissores e frutíferos da ciência mais recente são aqueles que, em vez de insistirem numa estéril contraposição entre fatores sociais, de um lado, e fatores biológicos, de outro, abraçam a noção de que o comportamento humano resulta de uma interação entre ambos. Com efeito, mesmo fenômenos sociais sofisticados e multifacetados são melhor compreendidos somando-se, ao lado dos olhares sociológico, antropológico, etc., as variáveis biológicas, presentes em sua gênese e desenvolvimento².

Do ponto de vista da ciência psicológica, décadas de estudos estabeleceram que o ato de punir não é resultado de um cálculo deliberado e racional, mas é, em grande medida, um processo cognitivo intuitivo e automático³. Violações de normas e agressões tendem a despertar uma sensação de ultraje moral nas vítimas e em testemunhas, e é essa reação instintiva que leva ao desejo de punir. O julgamento moral influencia a avaliação da gravidade de uma situação que o Direito Penal tipifica como crime⁴.

A Psicologia tradicionalmente realizou estudos utilizando métodos que perscrutavam a esfera comportamental, simulando situações em que sujeitos de pesquisa tinham que responder a questionários, ou que mediam reações a hipotéticas violações de normas. Nos últimos anos, porém, com a disponibilidade de técnicas e métodos de neuroimagem, tornou-se possível estudar não apenas o comportamento observável, mas também os padrões de funcionamento cerebral *in vivo* e em tempo real. Em virtude disso, tornou-se viável formular a questão de como são processadas as informações referentes a uma transgressão no cérebro.

Entender isso é a tarefa a que neurocientistas se propuseram nos últimos 15 anos, e assim uma profusão de experimentos se dedicou a delimitar estruturas neurais específicas que seriam recrutadas no cérebro do sujeito que toma a decisão de punir alguém. Segundo essa linha de estudos, tal como existem estruturas cerebrais muito especializadas no processamento de informações referentes à visão, à fala ou ao tato, existiria uma circuitaria cerebral específica para o ato de avaliar a culpabilidade e a necessidade de punir um transgressor.

Essa discussão é importante para a comunidade jurídica porque traz consequências para a forma como se encara a centralidade da punição no sistema normativo. Do ponto de vista das teorias da pena, a sanção penal historicamente foi justificada com argumentos preventivos ou retributivistas. Não se admite, contemporaneamente, que a vingança seja um argumento legítimo para a resposta penal a uma conduta delitiva; não obstante, o neoretributivismo é uma das tendências mais fortes na discussão teórica das últimas dé-

1 SAPOLSKY, Robert M. *Behave: the biology of humans at our best and worst*. New York: Penguin Press, 2017; Vide também, sobre o questionável (mas disseminado) argumento da exclusividade humana; WAAL, Frans de. *Are we smart enough to know how smart animals are?* New York: W. W. Norton & Company, 2016.

2 NETTLE, Daniel. *Tyneside neighbourhoods*. Cambridge, UK: Open Book Publishers, 2016.

3 CARLSMITH, Kevin M.; DARLEY, John M. Psychological aspects of retributive justice. *Advances in Experimental Social Psychology*, v. 40, n. 7, p. 193-236, 2008.

4 DARLEY, John M. Morality in the law: the psychological foundations of citizens' desires to punish transgressions. *Annual Review of Law and Social Science*, v. 5, n. 1, p. 1-23, 2009.

cadadas. Nas propostas mais recentes da dogmática penal, categorias conceituais como “reprovabilidade da conduta”, “censura” ou “denúnciação” figuram como finalidades da pena. Admite-se teoricamente, assim, que a gravidade da ofensa serve de medida para a proporcionalidade da pena – dito de outra forma, que o restabelecimento da justiça após uma transgressão passa pela sanção proporcional ao dano causado. Não raro, no debate público e acadêmico, sugere-se que há uma demanda psicológica das vítimas e dos cidadãos em geral por punição aos criminosos, oriunda de anseios retributivistas⁵.

As sanções jurídicas podem ser defendidas filosoficamente como maneiras de induzir a aderência a regras, especialmente por parte dos transgressores. Além dos mecanismos sociais de internalização de normas e de conformidade aos valores vigentes no grupo, o receio de sanções pode funcionar como um incentivo para que as normas vigentes no grupo sejam respeitadas⁶.

Na literatura que analisaremos a seguir, é frequente que se encontrem argumentos evolucionistas como explicações para a existência de uma circuitaria cerebral específica para o comportamento punitivo. Esse argumento evolucionista, que será melhor explicado no item seguinte, poderia ser sintetizado da seguinte maneira: num contexto interativo, em que os indivíduos podem ou não adotar um comportamento cooperativo, se os trapaceiros não são punidos, eles se locupletam, a confiança se perde, os grupos não conseguem mais cooperar, e a ordem social se desestabiliza. Grupos que punem os transgressores, por sua vez, funcionam de forma cooperativa e são bem-sucedidos. Nessa concepção, a punição dos transgressores é o que torna possível a vida em sociedade. Consequentemente, seria razoável supor que a evolução dos grupos humanos estaria intimamente associada à existência de uma determinada reação moral à transgressão, ou um mecanismo cognitivo para este fim.

A literatura em Psicologia Evolucionista frequentemente associa “módulos mentais” específicos a desafios encontrados pelos seres humanos no ambiente ancestral da espécie. As estruturas especializadas do cérebro humano, assim, seriam resultado de um processo de seleção natural ao longo de milhões de anos. Essa discussão, que é muito extensa para ser tratada aqui, está presente nos estudos em Neurociência da Punição, que geralmente associam a circuitaria cerebral engajada nas ações de avaliar uma transgressão, e de punir seu autor, ao desafio, presente no ambiente evolucionário, de manter a coesão dos grupos sociais⁷.

O que vale examinar, assim, é se, à luz do que se sabe atualmente na literatura especializada — e aqui estamos falando de experimentos recentes, publicados em periódicos científicos de prestígio, discutidos nas universidades de ponta —, é possível compreender o que *ocorre* no cérebro da pessoa que decide punir alguém que agiu contrariamente à norma ou à expectativa legítima de quem pune. E, se esses estudos nos dizem algo relevante, quais seriam as consequências disso para a discussão da área penal ou da política criminal?

Iniciaremos esse percurso estabelecendo os pressupostos teóricos da literatura em Neurociência da Punição. Em seguida, faremos uma revisão dos experimentos que vinculam comportamentos punitivos à atividade cerebral de áreas específicas. Por fim, analisaremos as limitações metodológicas e o problema das possíveis implicações dessa literatura para as políticas públicas da área criminal.

5 HORTA, Ricardo de Lins e. *A desalentadora função das prisões*: revisitando as teorias da pena à luz da psicologia experimental. Manuscrito.

6 STRUCHINER, Noel; CHRISMANN, Pedro H. V. Aspectos filosóficos e psicológicos das punições: Reunindo algumas peças do quebra-cabeça. *Caderno CRH*, v. 25, n. 2, p. 133-150, 2012; ALMEIDA, Fabio Portela Lopes. As origens evolutivas da cooperação humana, *Revista DireitoGV*, v. 9, n. 1, p. 243-268, 2013.

7 Aqui é relevante lembrar a forte influência que as perspectivas evolucionárias adquiriram na ciência psicológica *mainstream* nos últimos 40 anos. Embora a chamada Psicologia Evolucionista, como campo de investigação, ainda sofra diversas críticas, não há controvérsia científica relevante sobre a premissa de que o comportamento humano resulta de um processo de milhões de anos de pressões seletivas ao longo da evolução da espécie.

2. OS PRESSUPOSTOS TEÓRICOS DA NEUROCIÊNCIA DA PUNIÇÃO

Examinaremos, neste trabalho, os estudos que buscam experimentalmente entender como funciona o comportamento punitivo no cérebro humano. Antes, porém, é necessário analisar seus pressupostos. Os estudos de Neurociência da Punição partilham de duas hipóteses relevantes, uma de ordem evolucionista, outra de ordem neurocientífica: (i) a de que a cooperação humana foi estabilizada, em termos evolucionários, por meio do comportamento de punir trapaceiros, e (ii) o de que a punição, como comportamento antigo da espécie, recruta áreas cerebrais específicas, ligadas, respectivamente, ao estabelecimento da responsabilidade do agente e à determinação da punição que deveria ser aplicada no caso⁸.

É praticamente um truísmo na literatura psicológica mais recente afirmar que, no ambiente ancestral, grupos em que se puniam os aproveitadores cooperavam de forma mais estável, e, portanto, perseveraram. Desse modo, a circuitaria cerebral da punição cumpriria a função evolucionária essencial de dar efetividade às normas sociais, prevenindo que trapaceiros se beneficiem às custas dos outros.

O pano de fundo disso é a discussão do problema da evolução da cooperação na espécie humana, uma das questões que mais tem mobilizado estudiosos e estimulado debates entre biólogos, psicólogos, antropólogos e etólogos nos últimos 20 anos. A explicação a respeito da razão pela qual diversas espécies animais cooperam parece relativamente simples: animais não humanos, geralmente, cooperam apenas com indivíduos geneticamente aparentados, o que traz benefícios evolucionários evidentes, ao perpetuar seus genes. Algumas espécies também estabelecem relações de reciprocidade direta, em que os benefícios da interação são compartilhados com uma expectativa futura de retribuição.

Na espécie humana, porém, o nível de cooperação é ainda mais elevado. Em grupos humanos, a cooperação ocorre de forma relativamente comum entre indivíduos sem parentesco genético, e sem perspectiva de interação futura. Humanos ajudam estranhos, sem terem garantia ou certeza de que algum dia serão recompensados de volta por isso, apenas porque é o certo a se fazer. É a existência das normas morais que viabiliza, na espécie humana, a cooperação em larga escala entre indivíduos não aparentados, permitindo o surgimento e o desenvolvimento de grupos maiores, de cidades, de civilizações. O problema da evolução da cooperação é, assim, essencialmente, o problema de por que somos seres morais⁹.

Na narrativa mais disseminada, esse diferencial humano estaria ligado às normas sociais: os humanos são cooperadores condicionais, que estão dispostos a colaborar entre si se a regra vigente no grupo for a reciprocidade. Contudo, a presença de trapaceiros no grupo e a consequente falta de uma norma de reciprocidade fariam com que o comportamento não cooperativo, egoísta, fosse a melhor estratégia, e, portanto, a regra seria a não cooperação. Porém, se os trapaceiros são punidos, o comportamento aproveitador tem um custo, e assim se produz um estímulo para a aderência à norma de reciprocidade: se há sanção, a trapaça não compensa, e a norma social da reciprocidade vigora no grupo¹⁰.

Assim, esses estudos partem do pressuposto de que, para entender o comportamento punitivo, devemos antes sondar as origens evolucionárias da chamada “reciprocidade forte”: pessoas punem porque instintivamente é assim que se sentem obrigadas a fazer — não por acaso, punir um ofensor pode ser uma ação prazerosa, como de resto são outras que cumprem uma função evolucionária, como o sexo ou a saciedade alimentar. Nessa perspectiva, nutrir emoções negativas contra um aproveitador, e tomar a iniciativa de puni-lo, é uma ação pró-social: por essa razão, o termo “punição altruística”¹¹. Em termos evolucionários, incor-

8 KRUEGER, Frank; HOFFMAN, Morris. The emerging neuroscience of third-party punishment. *Trends in Neurosciences*, v. 39, n. 8, p. 499-501, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.tins.2016.06.004>>.

9 BAUMARD, Nicolas. *The origins of fairness: how evolution explains our moral nature*. New York: Oxford University Press, 2016.

10 FEHR, Ernst; FISCHBACHER, Urs. Social norms and human cooperation. *Trends in Cognitive Sciences*, v. 8, n. 4, p. 185-190, 2004.

11 FEHR, Ernst; GÄCHTER, Simon. Altruistic punishment in humans. *Nature*, v. 415, n. 6868, p. 137-140, 2002. Disponível em: <<http://www.nature.com/doi/10.1038/415137a>>.

rer em um custo para si mesmo ao punir um transgressor beneficia o grupo como um todo — correr o risco de se envolver num conflito com o transgressor é um sacrifício compensado pelo fato de que as normas serão cumpridas, e, assim, todos se beneficiarão de estar num grupo em que as normas sociais funcionam. Se não há punição, não há cooperação possível; se há punição, porém, todos são incentivados a cooperar. Assim, estaria resolvido o paradoxo: instintos punitivos estabilizam a cooperação¹².

É relevante compreender a distinção traçada por essa literatura entre as punições em primeira, segunda e terceira pessoa¹³. Ao causar mal a outrem, ao cometer uma agressão ou violação de norma social, o indivíduo poderia sofrer, como potencial consequência, três dimensões de “punição”. Em “primeira pessoa”, ela corresponde à culpa, remorso ou consciência interna que o próprio transgressor sentiria. Não sendo ela capaz de inibir o comportamento, poderia haver a “punição em segunda pessoa”, que pode ser caracterizada como “retaliação” — em que o próprio agredido retribuiria a ofensa por conta própria, num ato de legítima defesa ou vingança. E, por fim, a “punição em terceira pessoa” (*third-party punishment*), que consistiria no fato de um terceiro, não diretamente envolvido no conflito ou agressão, arriscar-se para atribuir ao transgressor a devida punição.

Esse papel de terceiro, que atualmente seria desempenhado de forma institucionalizada pelo sistema de justiça e de persecução penal, seria, na concepção dos defensores dessa narrativa da evolução dos comportamentos punitivos, correspondente ao chefe tribal, ou à figura dominante do grupo, em contextos ancestrais de caçadores-coletores. O terceiro, não originalmente envolvido no conflito, seria a figura neutra que faria cumprir a norma social vigente.

Desde já, esse modelo proposto requer explicar o seguinte problema: por que, do ponto de vista evolucionário, faria sentido se arriscar para punir um indivíduo que causou mal a outrem, e não a si mesmo? Punir alguém envolve o risco de reação, ou de se envolver num conflito originalmente surgido de conflitos de interesses não pertinentes a si mesmo. Logo, se é fácil buscar uma explicação evolucionária para a “punição de segunda pessoa” — retaliar um agressor é uma forma óbvia de cessar ou intimidar futuros atos de agressão — não é nada evidente a explicação da “punição em terceira pessoa”.

Buscando solucionar esse problema, entraram em cena os defensores das teorias da seleção multinível ou “seleção de grupo”. O terceiro que pune beneficia o grupo, ao manter a aplicabilidade da norma social, e, assim, indiretamente beneficia a si próprio. Grupos cooperativos tendem a prosperar, ao contrário de grupos onde não há cooperação — estes tendem a se extinguir, aqueles prosperam e deixam herdeiros. O comportamento punitivo seria, assim, um traço universal humano, que recrutaria circuitos neurais para uma função de extrema importância evolucionária¹⁴.

Esse modelo teórico evolucionário revelou-se altamente influente na literatura psicológica contemporânea, e onipresente na discussão sobre o comportamento de adesão a normas, sobre como ocorre a cooperação em sociedades, bem como a origem dos instintos morais humanos¹⁵. Não por acaso, seus proponentes e teorias são citados na introdução da maioria dos artigos científicos da Neurociência da Punição.

É preciso analisar esse ponto para compreender sua consequência lógica: se comportamentos humanos presentes desde o ambiente evolucionário ancestral acabam por recrutar áreas cerebrais bastante especia-

12 BOYD, R. et al. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, v. 100, n. 6, p. 3531-3535, 2003. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.0630443100>>.

13 HOFFMAN, Morris B. *The punisher's brain: the evolution of judge and jury*. New York: Cambridge University Press, 2014; KRUEGER, Frank; HOFFMAN, Morris. The emerging neuroscience of third-party punishment. *Trends in Neurosciences*, v. 39, n. 8, p. 499-501, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.tins.2016.06.004>>.

14 BUCKHOLTZ, Joshua W; MAROIS, René. The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience*, v. 15, n. 5, p. 655-661, 2012. Disponível em: <<http://www.nature.com/doi/10.1038/nn.3087>>.

15 BOWLES, Samuel; GINTIS, Herbert. *A cooperative species: human reciprocity and its evolution*. Princeton: Princeton University Press, 2011; HENRICH, Natalie; HENRICH, Joseph. *Why humans cooperate: a cultural and evolutionary explanation*. New York: Oxford University Press, 2007.

lizadas — e assim, o córtex occipital processa estímulos visuais, as áreas temporais do cérebro processam estímulos referentes à linguagem etc. —, então, o mesmo ocorreria com o comportamento punitivo. O cérebro teria um “módulo” especializado em punir transgressores.

Em seguida, revisaremos os estudos que giram em torno dessa narrativa, bem como as evidências neurocientíficas deles resultantes.

3. OS RESULTADOS DA NEUROCIÊNCIA DA PUNIÇÃO

Há *duas vertentes* de métodos experimentais adotados em Neurociência da Punição. Muitos deles proliferaram, nos últimos 15 anos, em paralelo com o crescimento de experimentos em Neuroeconomia, que combinam técnicas de neuroimagem com tarefas de interação social que envolvem ganhos e perdas num contexto de jogos econômicos. Assim, uma das vertentes desses estudos é o emprego de jogos como “Trust”, Ditador, Ultimato, Dilema dos Prisioneiros etc. Cria-se, num contexto de laboratório, uma situação de trapaça, seguida da possibilidade de punição do transgressor, e experimentalmente se utiliza de alguma técnica de neuroimagem. Assim, mede-se a atividade cerebral enquanto o sujeito reage ao fato de ser traçado ou de ver alguém ser vítima de trapaça.

Outra linhagem de pesquisas deriva da tradição dos estudos em Psicologia da Punição, em que textos curtos, vinhetas, são apresentadas aos sujeitos de pesquisa, geralmente descrevendo o cometimento de delitos ou de atos de transgressão, e os indivíduos devem assinalar um determinado nível de punição adequado ao caso¹⁶. Essa segunda vertente combina essas vinhetas fictícias com a neuroimagem.

Em ambas as linhas de estudos, verifica-se, desde então, o uso de técnicas neurocientíficas cada vez mais sofisticadas e precisas: os primeiros estudos utilizam o PET *scan*, além de ressonância magnética funcional (fMRI) e técnicas de estimulação transcraniana nos estudos mais recentes.

A literatura em Neurociência da Punição tem como um dos marcos inaugurais o experimento de cientistas da Universidade de Zurich que examina se a punição, ao ser um comportamento prazeroso, recrutaria áreas cerebrais ligadas ao sistema de recompensa¹⁷. Nesse estudo, foi conjugado um jogo econômico do estilo “*trust game*” com tomografia por emissão de pósitrons (PET *scan*) — um método comum à época, que investiga a ativação de áreas cerebrais valendo-se de contrastes radioativos.

No contexto do jogo, o sujeito recebia 10 unidades monetárias e podia escolher por dividi-la com outro jogador ou não. Se confiasse no outro jogador e repassasse a quantia, o montante era quadruplicado. Porém, o outro jogador podia tanto guardar tudo para si, trapaceando; ou dividir o montante multiplicado com o sujeito que havia confiado nele. Assim, o delineamento experimental simulava uma condição de trapaça, e, em seguida, avaliava cenários em que o sujeito podia punir ou não o trapaceiro, seja a custo zero, seja com custo em unidades monetárias. Os resultados revelaram que, ao punir o trapaceiro, de forma gratuita ou custosa, havia uma ativação no núcleo caudado, estrutura cerebral relacionada ao sistema de recompensa. A ativação dessa estrutura, aliás, estava correlacionada ao investimento na punição, no caso em que punir tinha custo. Além disso, o estudo mostrou ativação nas áreas pré-frontais mediais nos casos em que punir envolvia custos, em contraste com o cenário em que a punição era gratuita. Na interpretação dos pesquisadores, isso ocorreria porque punir com custos envolve um cálculo de custo-benefício: o investimento necessário no ato de punir *versus* o prazer derivado dele. Assim, o córtex pré-frontal integraria esse cálculo.

16 CARLSMITH, Kevin M.; DARLEY, John M. Psychological aspects of retributive justice. *Advances in Experimental Social Psychology*, v. 40, n. 7, p. 193-236, 2008.

17 QUERVAIN, D. J. F. de The neural basis of altruistic punishment. *Science*, v. 305, n. 5688, p. 1254-1258, 2004. Disponível em: <<http://www.sciencemag.org/cgi/doi/10.1126/science.1100735>>.

Importante destacar que esse experimento tratava da “punição em segunda pessoa”, ou retaliação, e não propriamente da punição por terceiros. A partir desse estudo inicial, outros se seguiram, focando em áreas de interesse próximas ao outrora chamado sistema límbico — tradicionalmente associado ao processamento da carga emocional, como é o caso da amígdala —, e as regiões pré-frontais — implicadas na literatura com comportamentos racionais e planejados. Com base em estudos com animais sobre estímulos aversivos, a amígdala, ligada ao medo e a reações aversivas, assim como a ínsula anterior, foram logo apontadas como áreas que poderiam estar ligadas ao processamento de informação nessas situações de trapaça. Áreas do córtex pré-frontal, por sua vez, estariam ligadas a comportamentos de cálculo de custo-benefício, ou de planejamento de cursos de ação possíveis, o que seria o caso da decisão de punir¹⁸.

Um dos primeiros experimentos a analisar a Neurociência da “punição em terceira pessoa”, e separando a etapa de determinação da culpabilidade da assinalação da punição, foi o de Joshua Buckholtz e colegas¹⁹. O método utilizado foi o de mostrar 50 vinhetas de crimes mais ou menos graves a 16 sujeitos submetidos à ressonância magnética funcional (fMRI). As condições experimentais eram: situações em que não havia cometimento de delitos; o cometimento de delitos leves, como furto; ou graves, como estupro e homicídio; e os mesmos crimes, mas com circunstâncias atenuantes da conduta do transgressor, que justificariam sua ação ou excluiriam a ilicitude da conduta. Os sujeitos tinham que, então, assinalar uma pena a cada situação, de uma escala de 0 (ausência de punição) a 9 (a punição mais severa possível).

Como resultado, o estudo mostrou associação entre a magnitude da punição assinalada e maior ativação da amígdala, córtex pré-frontal medial e córtex cingulado posterior, áreas ligadas ao processamento de estímulos socioemocionais. Para os autores, isso seria uma evidência do processamento emocional no momento de definição da punição. Por outro lado, verificou-se maior ativação da junção temporoparietal nos casos em que os indivíduos precisavam considerar as circunstâncias atenuantes da responsabilidade do ofensor — para os pesquisadores, essa área, ligada à Teoria da Mente, ou avaliação dos estados mentais alheios, seria recrutada para este fim. Finalmente, o córtex pré-frontal dorsolateral direito era mais ativado quando havia punição, tanto na presença de atenuantes, quanto na ausência delas.

Um ponto relevante é que os mesmos circuitos que vinham sendo implicados na “punição em segunda pessoa” seriam aqueles ativados nos casos em que um terceiro observador assinalaria a punição justa, na “punição em terceira pessoa”²⁰. A amígdala processaria a carga emocional do observador; em seguida, a integração entre o dano causado e a intenção do agente seria feita no córtex pré-frontal medial; e, posteriormente, o córtex pré-frontal dorsolateral sopesaria essas informações com aquelas oriundas de áreas posteriores (da junção temporoparietal e do sulco intraparietal) sobre o contexto e as intenções por trás do comportamento transgressor.

Em linha com esse modelo, um estudo posterior, com metodologia semelhante, buscou variar os cenários das vinhetas conforme sua carga emocional: o mesmo crime era descrito ou de forma relativamente neutra e objetiva, ou sensacionalista, com crueza nos detalhes. Novamente, verificava-se sempre, nos sinais da fMRI, ativação do córtex pré-frontal dorsolateral. Mas, mais importante, a ativação do *input* originário das amígdalas era fortemente aumentada nos cenários de descrição das vinhetas com componentes emocionais. Por sua vez, nos casos em que o dano causado era não intencional, o córtex cingulado dorsolateral e a junção temporoparietal entravam em cena, enviando sinais *top-down* para a amígdala e inibindo sua atividade. Dito de outra forma, se, por um lado, a amígdala enviava sinais para o córtex pré-frontal no caso de condutas graves; por outro, quando essas condutas eram escusáveis, um circuito temporoparietal-medial-pré-frontal suprimia a sua atividade²¹.

18 SEYMOUR, Ben; SINGER, Tania; DOLAN, Ray. The neurobiology of punishment. *Nature Reviews Neuroscience*, v. 8, n. 4, p. 300-311, 2007. Disponível em: <<http://www.nature.com/doi/10.1038/nrn2119>>.

19 BUCKHOLTZ, Joshua W. et al. The neural correlates of third-party Punishment. *Neuron*, v. 60, n. 5, p. 930-940, 2008.

20 BUCKHOLTZ; MAROIS, The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience*, v. 15, n. 5, p. 655-661, 2012. Disponível em: <<http://www.nature.com/doi/10.1038/nn.3087>>.

21 TREADWAY, Michael T. et al. Corticolimbic gating of emotion-driven punishment. *Nature Neuroscience*, v. 17, n. 9, p. 1270-1275, 2014. Disponível em: <<http://www.nature.com/neuro/journal/v17/n9/abs/nn.3781.html>>.

Nesse ponto, os pesquisadores consideravam ter alguma clareza sobre o papel da amígdala e dessa circuitaria temporoparietal, mas o papel do córtex dorsolateral persistia como um mistério. Sabia-se há tempos que o córtex pré-frontal está implicado nas funções executivas, na memória de trabalho, no planejamento de longo prazo, na integração superior de informações. Essa área, sempre, mostrava ativação em decisões de punir, mas não se sabia, exatamente, qual tipo de processamento ela estaria fazendo. A equipe de pesquisadores liderada por Joshua Buckholz, então, partiu para um outro método, a estimulação magnética transcraniana (TMS), capaz de desativar, seletivamente, áreas cerebrais com sinais magnéticos. Ao temporariamente “desligá-las”, seria em tese possível desvendar qual função desempenham.

Nesse experimento, 66 voluntários leram vinhetas de transgressões que, como nos estudos anteriores, descreviam ofensas das menos às mais graves, e que continham ou não circunstâncias atenuantes da responsabilidade do agente. Na condição de terceiros, precisavam decidir se alguma punição era devida, e qual seria, em cada um dos casos. Os resultados revelaram que, mesmo “desativando” o córtex pré-frontal dorsolateral, a avaliação da responsabilidade do agente não era afetada. Porém, a decisão de punir era afetada: com a aplicação da TMS, as punições assinaladas eram menos severas. O achado mais importante dessa pesquisa, assim, consistiu em mostrar que há uma dissociação entre o *juízo* a respeito da culpabilidade do agente e a integração de informações relevantes subjacentes à *decisão* sobre a intensidade da punição “justa”²².

Diante de objeções de que estudos com vinhetas e jogos econômicos seriam pouco realistas, outra equipe de pesquisadores se valeu de uma metodologia diferente. 33 participantes tinham que interagir com outros sujeitos enquanto estavam deitados no tubo de fMRI, onde poderiam levar choques ministrados por meio de agulhas intraepidérmicas, choques efetivamente dolorosos. Após receber os choques, os sujeitos eram informados a respeito do fato de haver ou não intenção, por parte de seu parceiro, de machucá-los. Além disso, os sujeitos eram informados se tinham, ou não, a opção de retaliar seu parceiro ou perdô-lo — um cenário que pode ser descrito como de “punição em segunda pessoa”, portanto. Como nem sempre a decisão de aplicar um choque era bem-sucedida, os pesquisadores puderam avaliar a condição de “tentativa” — intenção de agredir, frustrada por motivos alheios ao agente.

O estudo mostrou haver um circuito cerebral que ligaria a junção temporoparietal e o giro frontal inferior anterior (aIFG) à amígdala, inibindo a atividade desta quando o choque era acidental ou não intencional. Porém, se o outro participante tinha a intenção de ministrar o choque, mas falhava, a amígdala era estimulada pela ínsula anterior e pelo giro frontal inferior posterior (pIFG). Assim, o estudo mostraria que o processamento de informações sobre a intencionalidade do agente modularia a resposta emocional a uma agressão²³. Poderíamos, de forma ilustrativa, evocar o célebre dizer de Oliver Wendell Holmes: “até mesmo um cão sabe a diferença entre ser tropeçado e ser chutado”.

Outros achados convergentes vieram de um estudo de tomografia computadorizada feita com veteranos da Guerra do Vietnã que sofreram danos cerebrais²⁴. Estudos de lesão são tradicionais em Neuropsicologia, e ajudam a colocar à prova evidências oriundas de técnicas de neuroimagem. Nesse caso, os resultados mostraram que lesões nas áreas pré-frontais e parietais implicadas nos estudos anteriormente revisados afetavam a decisão de assinalar punição “em terceira pessoa” a condutas transgressoras.

Por fim, estudos mais recentes adotam técnicas estatísticas mais sofisticadas. Um estudo de vinhetas com 38 sujeitos e fMRI dissociou os momentos da decisão de punir por área cerebral: o processamento de informações sobre o dano causado (ínsula, regiões pré-frontais), sobre a intencionalidade do agente (circuitaria

22 BUCKHOLTZ, Joshua W. et al. From blame to punishment: disrupting prefrontal cortex activity reveals norm enforcement mechanisms. *Neuron*, v. 87, n. 6, p. 1369-1380, 2015.

23 YU, H.; LI, J.; ZHOU, X. Neural substrates of intention-consequence integration and its impact on reactive punishment in interpersonal transgression. *Journal of Neuroscience*, v. 35, n. 12, p. 4917-4925, 2015. Disponível em: <<http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3536-14.2015>>.

24 GLASS, Leila et al. Neural signatures of third-party punishment: evidence from penetrating traumatic brain injury. *Social Cognitive and Affective Neuroscience*, v. 11, n. 2, p. 253-262, 2015.

frontoparietal ligada à Teoria da Mente), e o momento em que se integra ambas informações (especialmente o córtex pré-frontal dorsolateral, com input das amígdalas). O modelo sugerido, porém, era ligeiramente mais complexo do que anteriormente proposto pelos estudos anteriores²⁵.

Em outro estudo recente com o uso de fMRI e vinhetas, focado na direcionalidade dos inputs dos circuitos, foi sugerido que o polo temporal e o córtex pré-frontal servem como “hubs” da “rede de mentalização”, a qual processa a informação sobre responsabilidade do agente, e, por sua vez, alimenta outras áreas como a junção temporoparietal e o córtex cingulado posterior²⁶.

Um outro estudo recente combinou um jogo chamado “Jogo da Justiça” com fMRI, visando comparar a ativação cerebral num contexto de “punição em segunda pessoa” com o de “punição em terceira pessoa”²⁷. Os sujeitos recebiam uma quantidade de recursos, estando ou na posição da vítima ou de terceiros-observadores. Então, um ofensor tomava parte desses recursos, e o sujeito de pesquisa podia gastar parte dos seus próprios recursos para puni-lo. Esse estudo previa, também, a hipótese de compensação — o sujeito de pesquisa que estivesse na posição de terceiro, vendo a vítima ser “furtada”, podia compensá-la com parte dos seus próprios recursos, ou punir o transgressor. Os resultados mostraram que, quanto maior o “furto”, mais os sujeitos queriam gastar seus próprios recursos para punir o transgressor — e que, na posição de terceiros, eles preferiam gastar para punir o transgressor a gastar para compensar a vítima. Outro achado relevante refere-se ao fato de que a quantidade de recursos gastos com a punição do transgressor era maior quando o sujeito era a própria vítima lesada, e não um terceiro observador.

Ainda como resultados, a decisão de punir recrutava uma área associada ao processamento de recompensas, o estriado ventral, ao passo que a decisão de compensar a vítima em vez de punir o transgressor revelava maior ativação da junção temporoparietal. Além disso, a “punição em segunda pessoa” envolvia maior ativação da ínsula anterior, ao passo que a “punição em terceira pessoa” mostrava maior ativação do córtex pré-frontal dorsolateral.

Em síntese, as mesmas áreas de interesse vêm sendo implicadas por diferentes estudos experimentais da chamada Neurociência da Punição. Numa formulação recente, foi proposto que os circuitos cerebrais envolvidos na decisão de punir seriam três, dissociáveis entre si, os quais processariam graus diferentes de informação para a decisão de punir um transgressor. Uma “rede de saliência” envolveria a ínsula anterior, o córtex cingulado anterior dorsal e a amígdala no processamento de informações aversivas, referentes à intensidade do *dano* causado pela transgressão. A segunda rede, ligada à “Teoria da Mente”, envolveria o córtex pré-frontal medial, o córtex cingulado posterior e a junção temporoparietal e integraria informações sobre o dano causado à vítima e a *intenção* do agente ofensor. Por fim, uma rede executiva central, ligada ao córtex pré-frontal dorsolateral, faria avaliações sobre as informações disponíveis e o contexto da ofensa, resultando numa *decisão* efetiva sobre a punição²⁸.

O grau de precisão desses estudos pode parecer surpreendente, e dar a impressão de que a “caixa preta” do comportamento de punir no cérebro está, definitivamente, sendo aberta. Todavia, embora impressionantes, esses resultados, ainda, dizem muito pouco, pois enfrentam limitações significativas. É o que analisaremos na sequência.

25 GINTHER, Matthew R. Parsing the behavioral and brain mechanisms of third-party punishment. *The Journal of Neuroscience*, v. 7, n. Sep, p. 9420-9434, 2016.

26 BELLUCCI, Gabriele et al. Effective connectivity of brain regions underlying third-party punishment: Functional MRI and Granger causality evidence. *Social Neuroscience*, v. 12, n. 2, p. 124-134, 2017. Disponível em: <<http://dx.doi.org/10.1080/17470919.2016.1153518>>.

27 STALLEN, Mirre et al. Neurobiological mechanisms of responding to injustice. *The Journal of Neuroscience*, v. 38, n. 12, p. 1242-1247, 2018. Disponível em: <<http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1242-17.2018>>.

28 KRUEGER, Frank; HOFFMAN, Morris. The emerging neuroscience of third-party punishment. *Trends in Neurosciences*, v. 39, n. 8, p. 499-501, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.tins.2016.06.004>>.

4. LIMITAÇÕES E PROBLEMAS DA NEUROCIÊNCIA DA PUNIÇÃO

As sofisticadas técnicas empregadas pela literatura de Neurociência da Punição mostram padrões de ativação de áreas cerebrais específicas em contextos experimentais em que se simulam injustiças ou transgressões. O que isso nos diz sobre a mente de quem decide punir? Quais inferências se pode extrair desses resultados? É importante, primeiramente, destacar que as interpretações desses dados dependem dos pressupostos teóricos adotados por essas pesquisas. E, ao menos até o presente momento, esses enfrentam objeções significativas.

Do ponto de vista metodológico, a ativação de áreas cerebrais específicas é um achado relevante na medida que confirma que o cérebro é um órgão altamente especializado, em que computações ocorrem de acordo com uma determinada estrutura citoarquitetônica, uma dinâmica de neurotransmissores e uma organização complexa de conexões neuronais. A base da Neurociência Cognitiva Social é o pressuposto de que o processamento de informações sociais ocorre em áreas específicas — as quais, se lesionadas, comprometidas ou afetadas, podem impactar o comportamento social do indivíduo. Porém, embora se saiba a função que várias dessas áreas desempenham, a dinâmica do “cérebro social” como um todo ainda é objeto de muita investigação²⁹.

Contudo, descrever um maior nível de oxigenação cerebral durante uma atividade, que é o que a maioria das técnicas de neuroimagem faz, não quer dizer que aquela área é exclusivamente ligada a um determinado comportamento. As estruturas cerebrais descritas não necessariamente são áreas cerebrais “para o fim de punição”. É possível arguir, por exemplo, que as áreas de interesse da Neurociência da Punição seriam estruturas de “domínio-geral”, isto é, ligadas ao comportamento interpessoal em geral, recrutadas para respostas comportamentais diversas no contexto social, e não apenas para a punição.

Um outro ponto de cautela relevante deriva da recente discussão sobre os limites metodológicos da ciência psicológica em geral. Russell Poldrack induziu a discussão no campo das neurociências a um novo patamar quando lembrou que verificar a ativação de áreas cerebrais *não* permite inferir estados mentais³⁰. Ademais, o método utilizado pelas técnicas de neuroimagem é o de subtração: comparam-se os padrões de funcionamento do cérebro em estado de repouso com aqueles padrões correlacionados ao momento de realização da tarefa experimental — ler uma vinheta, participar de um jogo econômico etc.

O que a neuroimagem “mede”, assim, não necessariamente é um *processo ou estado mental*; pode ser que ele aponte, apenas, estruturas relacionadas à adoção de uma determinada estratégia cognitiva, empregada para a solução da tarefa pelo sujeito de pesquisa. Em resumo: podemos medir padrões de ativação cerebral, mas segue sendo uma impropriedade metodológica sugerir, com base nessas técnicas, afirmações como “o sujeito sente prazer quando pune um ofensor”; ou ainda, “o sujeito sente indignação moral pela transgressão, como se pode inferir da ativação da amígdala”.

Além disso, uma objeção óbvia, mas importante de ser ressaltada, é o problema da validade ecológica de achados oriundos de contextos laboratoriais extremamente artificiais. Na discussão especializada, permanece aberto em debate a questão metodológica de até que ponto ser trapaceado num jogo de “*Trust*” ou “*Justiça*”, dentro de um tubo de ressonância magnética funcional, equivale à experiência de ser vítima ou testemunha de um delito, do ponto de vista da dinâmica cerebral. O que estudos com jogos econômicos mostram é o comportamento de pessoas... que participam de jogos econômicos. O contexto real de decisão é substancialmente mais complexo. É possível que os sujeitos de pesquisa estejam “pagando” para punir trapaceiros porque, num contexto de estudos num laboratório, em que as regras do jogo são explicitadas antes do experimento, eles

29 WANG, Yin; OLSON, Ingrid R. The original social network: white matter and social cognition. *Trends in Cognitive Sciences*, v. 22, n. 6, p. 1-13, 2018. Disponível em: <<https://doi.org/10.1016/j.tics.2018.03.005>>.

30 POLDRACK, Russell A. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, v. 72, n. 5, p. 692-697, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.neuron.2011.11.001>>.

apenas acreditem que o esperado deles é que optem pela opção de punir que lhes é disponibilizada. Pode ser que a existência de poucas alternativas, derivada da necessidade experimental de isolar variáveis, acabe induzindo comportamentos dos sujeitos — se só é possível punir ou não punir um trapaceiro, por que não o punir? É possível, ainda, que eles acreditem que é isso que os experimentadores esperem deles, ou a sinalização correta a se fazer. Não é possível inferir, com base nesse tipo de achado, porém, que exista um fenômeno universal humano chamado “punição altruísta” ou algo do gênero. É possível que o que os experimentos com jogos econômicos estejam demonstrando é que, em determinados contextos, as pessoas apostam recursos para revidar trapaceiros, e que as inferências amplas e profundas sobre o comportamento humano extraídas daí não passem de *artefatos* derivados da escolha metodológica dos pesquisadores³¹.

Para além das limitações metodológicas das ciências psicológicas em geral, porém, o estudo neurocientífico da punição enfrenta um desafio ainda mais formidável quando se examinam seus pressupostos teóricos. Não é pacífico na discussão em Psicologia Evolucionista que a “punição em terceira pessoa” seja um comportamento selecionado ao longo de milhões de anos.

A antropologia fornece importantes *insights* para a discussão de Psicologia Evolucionista, uma vez que a etnografia de sociedades tradicionais se debruça sobre grupos de caçadores-coletores relativamente pequenos — semelhantes à realidade que o *Homo sapiens* provavelmente experimentou na maior parte do seu tempo no planeta Terra. E, ao contrário do que sugerem os estudos da literatura revisada, as evidências antropológicas disponíveis em favor da existência da “punição de terceira pessoa” são escassas. Em sociedades tradicionais, geralmente, não há uma figura de autoridade neutra que pune transgressores — e é muito raro que alguém se arrisque a punir um ofensor, apenas, para defender “as normas vigentes no grupo”. Existe a chamada “punição em segunda pessoa”: se atacadas ou lesadas, as pessoas revidam. Mas ninguém incorre em custos para defender os interesses do grupo, como sugerido nos estudos de jogos econômicos³².

Uma alternativa mais vantajosa para indivíduos que se sentiram trapaceados é deixar de interagir com os ofensores. De fato, as evidências antropológicas apontam nesse sentido³³. Pessoas que contribuem com o grupo, que dividem recursos de forma equânime, que cooperam, têm boa reputação e são escolhidos com maior frequência para interações sociais. Já os aproveitadores sofrem desaprovação e deixam de ser vistos como sujeitos confiáveis. Assim, em comunidades tradicionais, a indiferença, o ostracismo, a popularmente dita “fofoca”, que visam destruir a reputação de um transgressor e isolá-lo no grupo, por exemplo, constituem formas de sanção frequentes³⁴. Embora, inequivocamente, elas sejam formas de sanção social, elas não se enquadram na definição que a literatura em Neurociência da Punição tem empregado para a “punição em terceira pessoa”: “incorrer em custo para si próprio para defender as regras vigentes”.

Não por acaso, a maior parte dos conflitos sociais não chega a ser objeto de intervenção da área penal. Seja em relações econômicas de mercado, seja em relações de amizade, as pessoas simplesmente expressam preferência por interagir com aquelas que lhes proporcionam as relações mais benéficas do seu ponto de vista. A maioria dos comportamentos desaprovados socialmente são objeto de sanções informais, e não necessariamente de punição. Humanos cooperam e agem moralmente de forma rotineira, não porque toda transgressão seja passível de punição, mas porque ser cooperativo é um comportamento que geralmente é reconhecido e recompensado pelos seus pares³⁵.

31 O trabalho mais completo ao apontar as diversas fraquezas metodológicas e impropriedades das inferências feitas a partir de experimentos com jogos econômicos é: GUALA, Francesco. Reciprocity: weak or strong? what punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, v. 35, n. 1, p. 1-15, 2012.

32 BAUMARD, Nicolas. Has punishment played a role in the evolution of cooperation?: a critical review. *Mind and Society*, p. 1-22, 2010.

33 FÜRER-HAIMENDORF, Chris Von. *Morals and merit: a study of values and social controls in South Asian Societies*. Chicago: The University of Chicago Press, 1967.

34 GURVEN, Michael. To give and to give not: the behavioral ecology of human food transfers. *Behavioral and Brain Sciences*, v. 27, n. 04, p. 543-583, 2004. Disponível em: <http://www.journals.cambridge.org/abstract_S0140525X04000123>.

35 GUALA, Francesco. Reciprocity: weak or strong? what punishment experiments do (and do not) demonstrate. *Behavioral and*

A própria tese evolucionária da “seleção de grupo” permanece sendo altamente controversa, havendo teóricos do campo que defendem que há outras explicações de origem darwinista para o problema da cooperação humana mais parcimoniosas e menos problemáticas³⁶. A evolução da moralidade humana pode ser explicada, também, por modelos mutualistas, como a Teoria da “*Partner Choice*”, que propõe que nossas intuições morais derivam da seleção de *indivíduos* mais cooperativos ao longo da história da espécie³⁷.

De fato, os experimentos mais recentes apontam que a busca pela reputação, ou de uma boa imagem social, é um motivador muito mais poderoso para a conduta moral que o medo de ser punido³⁸. Estudos utilizando novos paradigmas experimentais revelam que, diante de uma transgressão ou injustiça, pessoas se indignam e sentem raiva, mas não sacrificam a si próprios para punir quem lesou terceiros³⁹. Dito de outra forma, embora a “punição em segunda pessoa”, a retaliação, seja onipresente entre humanos — e mesmo em outras espécies⁴⁰ — a chamada “punição em terceira pessoa” pode ter tido sua existência exagerada por algumas pesquisas, e definitivamente não pode ser considerada a base da cooperação em sociedades humanas.

Toda essa literatura teórica e empírica, que contesta a tese da “punição altruísta” e da “seleção de grupo”, está ausente das revisões com as quais se iniciam os estudos de Neurociência da Punição.

De fato, a tese da “punição altruísta” sugere que quem pune transgressores, por se sacrificar pelo grupo em nome da vigência das normas, seria apreciado pela moralidade do seu comportamento. Porém, mesmo no paradigma de jogos econômicos, evidências mais recentes apontam que as pessoas atribuem melhor reputação e preferem mais interagir com quem ajuda vítimas, ou busca ampará-las, do que com quem toma a iniciativa de punir transgressores⁴¹.

Em suma, encarar a punição como única forma de restabelecer a balança da justiça acaba se revelando uma premissa teórica altamente contestável. A mente humana pode ter instintos vingativos ou retributivistas, mas não se resume a isso — ao contrário, se fosse esse o caso, não se verificariam em grupos sociais diversos comportamentos como as desculpas, o perdão e a reconciliação⁴².

Além disso, a tese da “punição altruísta” padece de uma tremenda pobreza antropológica — historicamente e ao redor do mundo, existem inúmeras respostas sociais possíveis a transgressões, que vão de mecanismos alternativos de composição de conflitos a formas institucionalizadas e bem-sucedidas de lidar com crimes graves, a exemplo das práticas de justiça restaurativa, um exemplo notável de política pública efetiva, baseada em evidências⁴³.

Brain Sciences, v. 35, n. 1, p. 1-15, 2012.

36 BAUMARD, Nicolas. *The origins of fairness: how evolution explains our moral nature*. New York: Oxford University Press, 2016; PINKER, Steven. The false allure of group selection. In: BUSS, David M. (Org.). *The handbook of evolutionary psychology*. Hoboken, New Jersey: John Wiley & Sons, 2016. p. 867-880.

37 BAUMARD, Nicolas; ANDRÉ, Jean-Baptiste; SPERBER, Dan. A mutualistic approach to morality: the evolution of fairness by partner choice. *Behavioral and Brain Sciences*, v. 36, p. 59-122, 2013.

38 GRIMALDA, Gianluca; PONDORFER, Andreas; TRACER, David P. Social image concerns promote cooperation more than altruistic punishment. *Nature Communications*, v. 7, p. 1-8, 2016. Disponível em: <<http://dx.doi.org/10.1038/ncomms12288>>.

39 PEDERSEN, Eric J.; MCAULIFFE, William H B; MCCULLOUGH, Michael E. The unresponsive avenger: more evidence that disinterested third parties do not punish altruistically. *Journal of Experimental Psychology: General*, v. 147, n. 4, p. 514-544, 2018. Disponível em: <[http://www.psy.miami.edu/faculty/mccullough/Papers/2018-Unresponsive Avenger.pdf](http://www.psy.miami.edu/faculty/mccullough/Papers/2018-Unresponsive%20Avenger.pdf)>.

PEDERSEN, E. J.; KURZBAN, R.; MCCULLOUGH, M. E. Do humans really punish altruistically? a closer look. *Proceedings of the Royal Society B: Biological Sciences*, v. 280, n. 1758, 2013. Disponível em: <<http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.2723>>.

40 RAIHANI, Nichola J.; THORNTON, Alex; BSHARY, Redouan. Punishment and cooperation in nature. *Trends in Ecology and Evolution*, v. 27, n. 5, p. 288-295, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.tree.2011.12.004>>.

41 PATIL, Indrajeet; DHALIWAL, Nathan A; CUSHMAN, Fiery. Reputational and cooperative benefits of third-party compensation. *PsyArXiv PrePrint*, p. 1-6, 2018.

42 MCCULLOUGH, Michael E.; KURZBAN, Robert; TABAK, Benjamin A. Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, v. 36, n. 01, p. 1-15, 2013. Disponível em: <http://www.journals.cambridge.org/abstract_S0140525X11002160>.

43 STRANG, Heather et al. Restorative justice conferencing (RJC) using face-to-face meetings of offenders and victims: effects

Por fim, e não menos importante, a decisão de punir alguém vai além de uma escolha num contexto experimental. No mundo real, a decisão de punir ocorre num contexto social dialógico, em que as pessoas analisam as consequências das ações alheias e produzem justificativas sobre por que punir — ou não. E, se por um lado, o retributivismo parece ser mesmo uma reação intuitiva automática a transgressões e injustiças, o processo de racionalização pode levar à revisão de decisões impulsivas⁴⁴. O esforço em justificar socialmente as próprias decisões exige que se produzam argumentos, e a entrada em cena do raciocínio deliberado pode reverter decisões intuitivas. Nesse sentido, a Neurociência da Punição se beneficiaria em incorporar, aos seus métodos experimentais, um fator sempre presente no contexto real da tomada de decisão humana: a necessidade de justificar as escolhas perante os pares⁴⁵.

5. CONSIDERAÇÕES FINAIS

As Neurociências têm avançado a passos largos para descrever o que ocorre, no nível neurofisiológico, quando as pessoas tomam decisões ou desempenham comportamentos complexos. Nesse sentido, é oportuno que a punição, fenômeno central no âmbito jurídico, seja objeto desse escrutínio.

Porém, uma revisão crítica da literatura disponível sugere que ainda é cedo para cantar vitória. A descrição de que certas estruturas cerebrais são recrutadas durante tarefas executadas em equipamentos de neuroimagem não permite que se façam inferências amplas sobre o papel do comportamento punitivo na mente humana, ou ainda, sobre sua suposta importância na trajetória evolucionária da espécie. Os pressupostos teóricos adotados ainda são muito mais frágeis do que alguns desses estudos parecem sugerir.

Entre os experimentos revisados e as estruturas sociais vigentes, há diversos níveis explicativos ainda em aberto. Quando um experimento neurocientífico, que conjuga neuroimagem com jogos econômicos, utiliza a palavra “punição”, refere-se a um fenômeno muito específico. Não se pode equiparar esse achado experimental com um conceito mais amplo de “sanção”, ou, mais ainda, com “pena”. Como mostramos neste artigo, há uma série de sanções informais, não penais, que são respostas rotineiras a comportamentos transgressivos e normas sociais. Além disso, para além da “retaliação”, da “vingança” e de outras respostas específicas, seres humanos também respondem a ofensas com comportamentos de reconciliação, compensação, reparação e perdão, e só recentemente estes têm sido incorporados por psicólogos nos desenhos experimentais.

Sobretudo, e aqui fica o alerta para os juristas que arriscam dar seus passos nesse campo interdisciplinar que vem sendo denominado “Neurodireito”: por mais interessantes que sejam os resultados obtidos até agora, eles são circunscritos aos contextos dos delineamentos experimentais de cada estudo. Não é possível afirmar, até o presente, que os seres humanos tenham uma “tendência inata”, um “instinto” ou “módulo mental específico” voltados para a punição de transgressores. O estado da arte não permite que se faça saltos teóricos entre estudos neurocientíficos e supostas justificativas de por que punimos na sociedade contemporânea. Especialmente por se tratar de um campo de pesquisa novo, que se embrenha em um tema ideologicamente carregado: o Direito Penal e as teorias da pena. Seria precipitado, assim, realizar inferências de política penal ou formular propostas de desenho institucional baseadas no que sabemos até agora em Neurociência da Punição.

on offender recidivism and victim satisfaction. *Campbell Systematic Reviews*, v. 12, 2013. Disponível em: <<https://campbellcollaboration.org/library/restorative-justice-conferencing-recidivism-victim-satisfaction.html>>.

44 AHARONI, Eyal; FRIDLUND, Alan J. Punishment without reason: isolating retribution in lay punishment of criminal offenders. *Psychology, Public Policy, and Law*, v. 18, n. 4, p. 599-625, 2012. Disponível em: <<http://doi.apa.org/getdoi.cfm?doi=10.1037/a0025821>>; FINCHER, Katrina M.; TETLOCK, Philip E. Brutality under cover of ambiguity. *Personality and Social Psychology Bulletin*, v. 41, n. 5, p. 629-642, 2015. Disponível em: <<http://journals.sagepub.com/doi/10.1177/0146167215571090>>.

45 MERCIER, Hugo; SPERBER, Dan. *The enigma of reason*. Cambridge, Massachusetts: Harvard University Press, 2017.

Esse alerta é importante porque, em obras recentes sobre o tema da Psicologia da Punição, especialmente no contexto estadunidense, tem sido sustentado, com base nos experimentos revisados, que as “intuições retributivistas” descobertas nos experimentos deveriam ser levadas em conta pelos formuladores da lei penal⁴⁶. Todavia, como pretendemos demonstrar neste trabalho, seria um erro afirmar que os resultados da Neurociência da Punição levam a tal conclusão. Muitos estudos ainda são necessários para que se tenha uma visão mais abrangente dos mecanismos psicológicos que entram em cena quando alguém é vítima ou testemunha de um delito. Somente a partir de uma teoria mais completa, seria possível pensar se as instituições punitivas que existem satisfazem os anseios de justiça das pessoas – e ainda assim, seriam necessários diversos argumentos morais, políticos e de política pública para um debate de uma questão dessa magnitude.

Um último ponto merece destaque ao se tratar dos estudos experimentais sobre a punição e o que eles podem acrescentar ao debate teórico penal. Do ponto de vista psicológico, pode-se afirmar que a punição funciona? A análise comportamental já nos respondeu essa pergunta há muito tempo: sim, mas em circunstâncias muito específicas — desde que haja regras muito claras, que ela ocorra sempre e imediatamente depois que um comportamento indesejável ocorre⁴⁷. O sistema punitivo que temos está muito longe de cumprir esses requisitos. A sanção penal, quando ocorre, é errática, seletiva, morosa — nem todos os que cometem delitos são descobertos, nem todos os delitos são investigados, nem todos os acusados chegam a ser sentenciados; injustiças ocorrem, inocentes são punidos, ofensores podem sair impunes, a depender da capacidade investigativa dos órgãos de persecução penal, da seletividade do sistema, etc. Em suma, a reação jurídica a delitos vigente não poderia ser mais contrária ao que as ciências comportamentais preconizam que “funciona”. Não por acaso, a pena de prisão, essa recentíssima invenção humana, não apenas tem efeitos preventivos altamente contestáveis, como há evidências significativas de seu caráter criminogênico⁴⁸.

Aí reside o maior risco enfrentado atualmente pela chamada Neurociência da Punição: ela dá muita centralidade a um fenômeno que não é a única forma que a mente humana tem de reagir a transgressões, podendo até mesmo se prestar a justificar respostas jurídicas empiricamente ineficazes. Caracterizar a punição como forma “natural” de reação pode ser uma forma de legitimar um sistema falido e, pior, impedir que métodos alternativos de lidar com conflitos ganhem espaço. Contra esse risco, insistimos que a boa ciência aplicada não é aquela que limita; é aquela que *amplia* as possibilidades de imaginação institucional.

REFERÊNCIA

AHARONI, Eyal; FRIDLUND, Alan J. Punishment without reason: isolating retribution in lay punishment of criminal offenders. *Psychology, Public Policy, and Law*, v. 18, n. 4, p. 599-625, 2012. Disponível em: <<http://doi.apa.org/getdoi.cfm?doi=10.1037/a0025821>>.

ALMEIDA, Fabio Portela Lopes. As origens evolutivas da cooperação humana. *Revista DireitoGV*, v. 9, n. 1, p. 243-268, 2013.

BAUMARD, Nicolas. Has punishment played a role in the evolution of cooperation?: a critical review. *Mind and Society*, p. 1-22, 2010.

BAUMARD, Nicolas. *The origins of fairness: how evolution explains our moral nature*. New York: Oxford University Press, 2016.

46 Um dos expoentes dessa tendência recente é Paul H. Robinson, renomado pesquisador que sugere que as “intuições de justiça” partilhadas pela comunidade deveriam ser incorporadas à lei penal como forma de fortalecer sua legitimidade social. Implícita nessa proposta está uma opção teórica pelo neorretributivismo como justificativa legítima para a lei penal.

47 MILTENBERGER, Raymond G. *Behavior modification: principles & procedures*. 5. ed. Belmont, CA: Cengage Learning, 2012. p. 101-119.

48 NAGIN, Daniel S. Deterrence in the twenty-first century. *Crime and Justice*, v. 42, n. 1, p. 199-263, 2013.

- BAUMARD, Nicolas; ANDRÉ, Jean-Baptiste; SPERBER, Dan. A mutualistic approach to morality: the evolution of fairness by partner choice. *Behavioral and Brain Sciences*, v. 36, p. 59-122, 2013.
- BELLUCCI, Gabriele et al. Effective connectivity of brain regions underlying third-party punishment: Functional MRI and Granger causality evidence. *Social Neuroscience*, v. 12, n. 2, p. 124-134, 2017. Disponível em: <<http://dx.doi.org/10.1080/17470919.2016.1153518>>.
- BOWLES, Samuel; GINTIS, Herbert. *A cooperative species: human reciprocity and its evolution*. Princeton: Princeton University Press, 2011.
- BOYD, R. et al. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, v. 100, n. 6, p. 3531-3535, 2003. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.0630443100>>.
- BUCKHOLTZ, Joshua W. et al. From blame to punishment: disrupting prefrontal cortex activity reveals norm enforcement mechanisms. *Neuron*, v. 87, n. 6, p. 1369-1380, 2015.
- BUCKHOLTZ, Joshua W. et al. The neural correlates of third-party Punishment. *Neuron*, v. 60, n. 5, p. 930-940, 2008.
- BUCKHOLTZ, Joshua W.; MAROIS, René. The roots of modern justice: cognitive and neural foundations of social norms and their enforcement. *Nature Neuroscience*, v. 15, n. 5, p. 655-661, 2012. Disponível em: <<http://www.nature.com/doi/10.1038/nn.3087>>.
- CARLSMITH, Kevin M.; DARLEY, John M. Psychological aspects of retributive justice. *Advances in Experimental Social Psychology*, v. 40, n. 7, p. 193-236, 2008.
- DARLEY, John M. Morality in the law: the psychological foundations of citizens' desires to punish transgressions. *Annual Review of Law and Social Science*, v. 5, n. 1, p. 1-23, 2009.
- FEHR, Ernst; FISCHBACHER, Urs. Social norms and human cooperation. *Trends in Cognitive Sciences*, v. 8, n. 4, p. 185-190, 2004.
- FEHR, Ernst; GÄCHTER, Simon. Altruistic punishment in humans. *Nature*, v. 415, n. 6868, p. 137-140, 2002. Disponível em: <<http://www.nature.com/doi/10.1038/415137a>>.
- FINCHER, Katrina M.; TETLOCK, Philip E. Brutality under cover of ambiguity. *Personality and Social Psychology Bulletin*, v. 41, n. 5, p. 629-642, 2015. Disponível em: <<http://journals.sagepub.com/doi/10.1177/0146167215571090>>.
- FÜRER-HAIMENDORF, Chris Von. *Morals and merit: a study of values and social controls in South Asian Societies*. Chicago: The University of Chicago Press, 1967.
- GINTHER, Matthew R. Parsing the behavioral and brain mechanisms of third-party punishment. *The Journal of Neuroscience*, v. 7, n. Sep, p. 9420-9434, 2016.
- GLASS, Leila et al. Neural signatures of third-party punishment: evidence from penetrating traumatic brain injury. *Social Cognitive and Affective Neuroscience*, v. 11, n. 2, p. 253-262, 2015.
- GRIMALDA, Gianluca; PONDORFER, Andreas; TRACER, David P. Social image concerns promote cooperation more than altruistic punishment. *Nature Communications*, v. 7, p. 1-8, 2016. Disponível em: <<http://dx.doi.org/10.1038/ncomms12288>>.
- GUALA, Francesco. Reciprocity: weak or strong? what punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, v. 35, n. 1, p. 1-15, 2012.
- GURVEN, Michael. To give and to give not: the behavioral ecology of human food transfers. *Behavioral and Brain Sciences*, v. 27, n. 04, p. 543-583, 2004. Disponível em: <http://www.journals.cambridge.org/abstract_S0140525X04000123>.

HENRICH, Natalie; HENRICH, Joseph. *Why humans cooperate: a cultural and evolutionary explanation*. New York: Oxford University Press, 2007.

HOFFMAN, Morris B. *The punisher's brain: the evolution of judge and jury*. New York: Cambridge University Press, 2014.

HORTA, Ricardo de Lins e. *A desalentadora função das prisões: revisitando as teorias da pena à luz da psicologia experimental*. Manuscrito.

KRUEGER, Frank; HOFFMAN, Morris. The emerging neuroscience of third-party punishment. *Trends in Neurosciences*, v. 39, n. 8, p. 499-501, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.tins.2016.06.004>>.

MCCULLOUGH, Michael E.; KURZBAN, Robert; TABAK, Benjamin A. Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, v. 36, n. 01, p. 1-15, 2013. Disponível em: <http://www.journals.cambridge.org/abstract_S0140525X11002160>.

MERCIER, Hugo; SPERBER, Dan. *The enigma of reason*. Cambridge, Massachusetts: Harvard University Press, 2017.

MILTENBERGER, Raymond G. *Behavior modification: principles & procedures*. 5. ed. Belmont, CA: Cengage Learning, 2012.

NAGIN, Daniel S. Deterrence in the twenty-first century. *Crime and Justice*, v. 42, n. 1, p. 199-263, 2013.

NETTLE, Daniel. *Tyneside neighbourhoods*. Cambridge, UK: Open Book Publishers, 2016.

PATIL, Indrajeet; DHALIWAL, Nathan A.; CUSHMAN, Fiery. Reputational and cooperative benefits of third-party compensation. *PsyArXiv PrePrint*, p. 1-56, 2018.

PEDERSEN, E. J.; KURZBAN, R.; MCCULLOUGH, M. E. Do humans really punish altruistically? a closer look. *Proceedings of the Royal Society B: Biological Sciences*, v. 280, n. 1758, 2013. Disponível em: <<http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2012.2723>>.

PEDERSEN, Eric J.; MCAULIFFE, William H. B.; MCCULLOUGH, Michael E. The unresponsive avenger: more evidence that disinterested third parties do not punish altruistically. *Journal of Experimental Psychology: General*, v. 147, n. 4, p. 514-544, 2018. Disponível em: <[http://www.psy.miami.edu/faculty/mmccullough/Papers/2018-Unresponsive Avenger.pdf](http://www.psy.miami.edu/faculty/mmccullough/Papers/2018-Unresponsive%20Avenger.pdf)>.

PINKER, Steven. The false allure of group selection. In: BUSS, David M. (Org.). *The handbook of evolutionary psychology*. Hoboken, New Jersey: John Wiley & Sons, 2016. p. 867-880.

POLDRACK, Russell A. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, v. 72, n. 5, p. 692-697, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.neuron.2011.11.001>>.

QUERVAIN, D. J. F. The neural basis of altruistic punishment. *Science*, v. 305, n. 5688, p. 1254-1258, 2004. Disponível em: <<http://www.sciencemag.org/cgi/doi/10.1126/science.1100735>>.

RAIHANI, Nichola J.; THORNTON, Alex; BSHARY, Redouan. Punishment and cooperation in nature. *Trends in Ecology and Evolution*, v. 27, n. 5, p. 288-295, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.tree.2011.12.004>>.

SAPOLSKY, Robert M. *Behave: the biology of humans at our best and worst*. New York: Penguin Press, 2017.

SEYMOUR, Ben; SINGER, Tania; DOLAN, Ray. The neurobiology of punishment. *Nature Reviews Neuroscience*, v. 8, n. 4, p. 300-311, 2007. Disponível em: <<http://www.nature.com/doi/10.1038/nrn2119>>.

STALLEN, Mirre et al. Neurobiological mechanisms of responding to injustice. *The Journal of Neuroscience*,

v. 38, n. 12, p. 1242-1247, 2018. Disponível em: <<http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1242-17.2018>>.

STRANG, Heather et al. Restorative justice conferencing (RJC) using face-to-face meetings of offenders and victims: effects on offender recidivism and victim satisfaction. *Campbell Systematic Reviews*, v. 12, 2013. Disponível em: <<https://campbellcollaboration.org/library/restorative-justice-conferencing-recidivism-victim-satisfaction.html>>.

STRUCHINER, Noel; CHRISMANN, Pedro H. V. Aspectos filosóficos e psicológicos das punições: Reunindo algumas peças do quebra-cabeça. *Caderno CRH*, v. 25, n. 2, p. 133-150, 2012.

TREADWAY, Michael T. et al. Corticolimbic gating of emotion-driven punishment. *Nature Neuroscience*, v. 17, n. 9, p. 1270-1275, 2014. Disponível em: <<http://www.nature.com/neuro/journal/v17/n9/abs/nn.3781.html>>.

WAAL, Frans de. *Are we smart enough to know how smart animals are?* New York: W. W. Norton & Company, 2016.

WANG, Yin; OLSON, Ingrid R. The original social network: white matter and social cognition. *Trends in Cognitive Sciences*, v. 22, n. 6, p. 1-13, 2018. Disponível em: <<https://doi.org/10.1016/j.tics.2018.03.005>>.

YU, H.; LI, J.; ZHOU, X. Neural substrates of intention-consequence integration and its impact on reactive punishment in interpersonal transgression. *Journal of Neuroscience*, v. 35, n. 12, p. 4917-4925, 2015. Disponível em: <<http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3536-14.2015>>.

Para publicar na revista Brasileira de Políticas Públicas, acesse o endereço eletrônico www.rbpp.uniceub.br
Observe as normas de publicação, para facilitar e agilizar o trabalho de edição.