

RAZIONALITÀ, CAPACITÀ, IDENTITÀ

Nuove frontiere della responsabilità giuridica



a cura di

Antonio Carratta

Mario De Caro

Giorgio Pino



Università degli Studi Roma Tre
Dipartimento di Giurisprudenza

NELLA STESSA COLLANA

1. A. CARRATTA, M. DE CARO, G. PINO (a cura di), *Intorno al ragionamento giuridico*, 2022

RAZIONALITÀ, CAPACITÀ, IDENTITÀ

Nuove frontiere della responsabilità giuridica

a cura di

Antonio Carratta

Mario De Caro

Giorgio Pino



Roma TrE-Press

2023

Direttori della Collana Dentro il giudizio

Antonio Carratta, *Università Roma Tre*

Mario De Caro, *Università Roma Tre*

Giorgio Pino, *Università Roma Tre*

Collana pubblicata nel rispetto del Codice etico adottato dal Dipartimento di Giurisprudenza dell'Università degli Studi Roma Tre, in data 22 aprile 2020.

Coordinamento editoriale:

Gruppo di Lavoro *RomaTrE-Press*

Elaborazione grafica della copertina: **MOSQUITO**, mosquitoroma.it

Caratteri tipografici utilizzati:

Abadi MT, Aceh Light, Barlow Semi Condensed, Pluto Sans W01 (copertina e frontespizio)

Adobe Garamond Pro (testo)

Impaginazione e cura editoriale: Colitti-Roma colitti.it

Edizioni: RomaTrE-Press ©

Roma, febbraio 2023

ISBN: 979-12-5977-153-7

<http://romatrepress.uniroma3.it>

Questa opera è assoggettata alla disciplina *Creative Commons attribution 4.0 International License* (CC BY-NC-ND 4.0) che impone l'attribuzione della paternità dell'opera, proibisce di alterarla, trasformarla o usarla per produrre un'altra opera, e ne esclude l'uso per ricavarne un profitto commerciale.



L'attività della *RomaTrE-Press* è svolta nell'ambito della

Fondazione Roma Tre-Education, piazza della Repubblica 10, 00185 Roma

Dentro il giudizio
Dialoghi di Diritto e Filosofia

“Giudizio”, così come “ragionamento giuridico”, è il nome che conviene ad un insieme, alquanto eterogeneo, di operazioni che accompagna l’applicazione delle norme giuridiche alla realtà. La vita del diritto, delle norme giuridiche, è un movimento dialettico tra ‘produzione’ e ‘applicazione’: le norme nascono come entità generali e astratte, destinate a regolare serie indefinite di casi, ma poi devono necessariamente essere calate in vicende della vita, nella decisione di controversie e casi concreti. Il percorso dalla produzione all’applicazione delle norme è raramente lineare, spesso faticoso e laborioso, e talvolta anche circolare – le vicende applicative spesso suggeriscono agli organi della produzione giuridica la necessità di rivedere le norme, di riformularle, quando non addirittura di cancellarle per via di abrogazione o dichiarazione di incostituzionalità.

Nel giudizio convergono sapienza tecnico-giuridica, valutazioni morali, senso comune, saperi specialistici, considerazioni strategiche, conformistiche e opportunistiche, intuizioni e percorsi psicologici talvolta opachi anche a chi li pone in essere. Nonostante gli studi sul ragionamento giuridico abbiano assunto un ruolo di primo piano da almeno cinquant’anni, una vasta *terra incognita* rimane ancora da esplorare, da ultimo grazie anche ai progressi vertiginosi nel campo delle scienze cognitive e dell’intelligenza artificiale.

La collana *Dentro il giudizio* si propone di raccogliere, sollecitare e diffondere in modalità accesso aperto studi e ricerche sul giudizio e sul ragionamento giuridico, facendo dialogare il diritto con saperi quali la filosofia, la psicologia, l’epistemologia, l’economia, la sociologia, l’antropologia, e incoraggiando approcci innovativi ai problemi del ragionamento giuridico, non solo nella prospettiva del dibattito specialistico ma anche al fine di rendere sempre più accessibili queste tematiche a fini didattici.

ANTONIO CARRATTA
MARIO DE CARO
GIORGIO PINO

Indice

<i>Presentazione</i> , di ANTONIO CARRATTA, MARIO DE CARO, GIORGIO PINO	9
MASSIMO REICHLIN, <i>Emozioni e processi riflessivi nei giudizi morali</i>	11
1. <i>Razionalismo e sentimentalismo nella tradizione filosofica</i>	11
2. <i>Neosentimentalismo</i>	13
3. <i>La teoria del doppio processo</i>	18
4. <i>Aspetti critici e approcci alternativi</i>	23
5. <i>Conclusioni</i>	27
FLAVIA CORSO, ANDREA LAVAZZA, <i>Diritto e neuroscienze: un avvicinamento a due velocità</i>	29
1. <i>Premessa</i>	29
2. <i>Neurodiritto, definizione e contesto</i>	30
2.1. <i>La prospettiva neuroscientifica</i>	33
2.2. <i>“Trattare” il crimine</i>	35
3. <i>La realtà giuridica e del senso comune</i>	37
3.1. <i>Il compatibilismo</i>	39
3.2. <i>La tesi antinaturalistica</i>	41
3.3. <i>Il ruolo delle emozioni</i>	43
4. <i>Il situazionismo</i>	45
5. <i>Conclusione</i>	49
MARIO DE CARO, <i>Retributivismo, utilitarismo e libero arbitrio</i>	51
1. <i>Utilitarismo radicale in teoria della pena</i>	51
2. <i>La sfida deterministica</i>	54
3. <i>La sfida dell'epifenomenismo</i>	59
4. <i>La teoria della pena tra utilitarismo e retributivismo</i>	61
LORENZO MILAZZO, <i>Libero arbitrio e colpevolezza nell'ordinamento giuridico italiano</i>	67
1. <i>Una premessa</i>	67
2. <i>Il Codice Rocco e i lavori preparatori</i>	69
3. <i>Principio di colpevolezza e libero arbitrio nell'ordinamento repubblicano</i>	75
4. <i>Una prima difficoltà, d'ordine logico</i>	80
5. <i>Dalla libertà del volere alla «motivabilità mediante norme»</i>	82
6. <i>Dalla possibilità di assumere l'atteggiamento doveroso della volontà alla constatazione del fatto che gli altri in genere lo fanno</i>	86
7. <i>'Come se'</i>	89

DANIELE PIVA, <i>Il ruolo delle emozioni nel giudizio penale tra “oggetto della valutazione” e “valutazione dell’oggetto”</i>	95
SOFIA BONICALZI, <i>Scatole nere, pregiudizi e discrezionalità. Tre questioni etiche su intelligenza artificiale e diritto</i>	103
1. <i>Introduzione</i>	103
2. <i>Il problema della scatola nera</i>	105
3. <i>Pregiudizi umani e algoritmici</i>	110
4. <i>Giustizia distributiva e retributiva fra discrezionalità e uniformità</i>	114
5. <i>Conclusioni</i>	121
MARTA INFANTINO, <i>Responsabilità da danno algoritmico</i>	123
1. <i>Introduzione</i>	123
2. <i>Prospettive e caveat</i>	125
3. <i>Lo spettro degli incidenti algoritmici</i>	128
4. <i>Connotati e problemi</i>	131
5. <i>Illeciti algoritmici e processi extracontrattuali</i>	134
5.1. <i>Giudici e giurie</i>	134
5.2. <i>Profili probatori</i>	135
5.3. <i>Danni di massa e azioni collettive</i>	136
5.4. <i>Il costo delle liti</i>	137
6. <i>Conclusioni</i>	139

*Presentazione**

ANTONIO CARRATTA, MARIO DE CARO, GIORGIO PINO

Il diritto della modernità, dall'età delle codificazioni in avanti, si è sviluppato assumendo una certa idea di soggetto di diritto, di *homo juridicus*: un soggetto razionale, pienamente capace di intendere e di volere e dunque "responsabile" per i propri atti, leciti e illeciti – salva la dimostrazione dell'assenza, intesa come fatto eccezionale, come eccezione alla regola, di tale piena capacità. L'*homo juridicus* è un soggetto che sa ciò che vuole, e che è in grado di operare un attento calcolo costi-benefici tra un certo corso di azione e le conseguenze ad esso associate dall'ordinamento giuridico. E tali capacità razionali, in fondo, sono ciò che ci rende tutti uguali davanti alla legge – tutti, parimenti, soggetti di diritto. Il soggetto di diritto è un soggetto (apparentemente) astratto – asessuato, non situato in specifiche condizioni sociali, personali, economiche: è semplicemente un soggetto capace di autodeterminarsi, e conseguentemente di incidere nel mondo delle relazioni giuridiche sia costituendo rapporti giuridicamente rilevanti, sia potendo essere chiamato a rispondere per le conseguenze delle proprie determinazioni.

Ora, è noto come questa raffigurazione del soggetto di diritto, che è in effetti una delle "grandi narrazioni", o delle "mitologie giuridiche", della modernità, sia stata progressivamente smantellata, decostruita, nel corso della seconda metà del Novecento. Per un verso, quella del soggetto "astratto" di diritto è stata chiaramente riconosciuta come una maschera ideologica, che lo stesso diritto positivo ha provveduto progressivamente a lacerare: basti vedere come negli articoli della nostra Costituzione il soggetto "unico" e "astratto" di diritto lasci il posto ad una vasta serie di "soggetti situati" colti nelle loro specifiche condizioni e relazioni sociali: il lavoratore, la lavoratrice, i figli, lo studente, i capaci e meritevoli ma privi di mezzi, i soggetti bisognosi di cure...

Per altro verso, e più recentemente, ad andare progressivamente in crisi è stata proprio l'idea del soggetto di diritto come soggetto razionale e

* Un sentito ringraziamento va alla dott.ssa Noemi Calidori per l'attenta cura redazionale svolta sui testi.

dunque capace di “autonomia” e di “responsabilità”, quantomeno davanti al diritto. Le scienze cognitive e le neuroscienze stanno facendo emergere un nuovo paradigma secondo il quale gran parte delle cose che accadono nel nostro cervello restano a noi stessi imperscrutabili: molti dei nostri processi decisionali, forse tutti, sfuggono al nostro controllo o comunque non sappiamo esattamente perché li abbiamo adottati. E inoltre, si affacciano sempre di più decisioni prese da soggetti non umani (e in particolare da applicazioni di intelligenza artificiale), attraverso processi a loro volta opachi e sostanzialmente imprevedibili – decisioni che diventano sempre più rilevanti per il modo del diritto, e per i meccanismi giuridici di attribuzione di responsabilità per fatti dannosi (si pensi al caso di danni causati da veicoli a guida autonoma).

È a questioni di questo tipo che è stato dedicato il secondo ciclo di seminari “Dentro il giudizio. Dialoghi di diritto e filosofia” (Università Roma Tre, primavera 2022), nell’ambito del quale cui sono stati originariamente presentati i saggi ora raccolti in questo volume.

I contributi esplorano, nella prospettiva di una riflessione multidisciplinare, vari aspetti problematici attinenti al superamento del paradigma del soggetto razionale: il rapporto tra sentimenti, emozioni e razionalità nei giudizi morali (M. Reichlin); il ricorso alle neuroscienze per esplorare la nozione di “libera determinazione” da parte del soggetto agente (F. Corso, A. Lavazza); la problematica configurabilità del concetto di libero arbitrio, e le conseguenze di ciò per l’imputabilità della responsabilità penale (M. De Caro, L. Milazzo); il ruolo delle emozioni tanto nella realizzazione della condotta penalmente rilevante, quanto nella valutazione che di tale condotta è chiamato a fare il giudice (D. Piva); le questioni etiche e giuridiche associate alle decisioni prese da un’intelligenza artificiale (S. Bonicalzi, M. Infantino).

I problemi evocati sono enormi e tanto complessi che si fa fatica a coglierne tutte le implicazioni, il panorama tecnologico e la ricerca scientifica ancora in piena evoluzione, le soluzioni solo provvisorie. Ci auguriamo che questo volume possa apportare un utile contributo ad una discussione che assume una sempre maggiore centralità nelle vite quotidiane di tutti, così come nella riflessione filosofica e nella pratica giuridica, e che – davanti a inediti rischi per i diritti della persona – oggi più che mai chiama il giurista ad un’attenta opera progettuale di soluzioni innovative.

Emozioni e processi riflessivi nei giudizi morali

MASSIMO REICHLIN*

SOMMARIO: 1. Razionalismo e sentimentalismo nella tradizione filosofica – 2. Neosentimentalismo – 3. La teoria del doppio processo – 4. Aspetti critici e approcci alternativi – 5. Conclusioni.

1. Razionalismo e sentimentalismo nella tradizione filosofica

Larga parte della tradizione filosofica occidentale concepisce la morale in termini che potremmo definire razionalisti. Ciò indica almeno due cose: da un lato, la tesi secondo cui i giudizi morali sono opera della ragione, dall'altro l'idea che la ragione abbia carattere pratico, ovvero che sia in grado di fornire ragioni o motivi per compiere certe azioni e per astenersi da altre. Con tinte più o meno decise, questa concezione razionalista attraversa l'intera etica occidentale, da Platone e Aristotele fino a Tommaso d'Aquino e al razionalismo moderno, trovando espressione paradigmatica nell'etica di Kant. Quest'ultimo, come è noto, concepisce la legge morale come espressione della *pura* ragione, dove la purezza razionale si realizza escludendo qualunque elemento sensibile, legato alle inclinazioni, da ciò che determina la volontà e quindi l'azione; le inclinazioni sono per Kant intrinsecamente passive, ovvero qualcosa che semplicemente *accade* all'essere umano e che pertanto non coinvolge la sua libertà; al contrario, l'uso della ragione ha carattere attivo ed esprime la sua libera determinazione. Secondo Kant, «La pura ragione è per sé sola pratica, e dà (all'uomo) una legge universale, che chiamiamo *legge morale*»¹. Tale legge coincide, in realtà, con l'*autonomia* della stessa ragion pura pratica, ossia con la sua capacità di determinare la volontà umana senza presupporre incentivi di carattere empirico; ogni elemento empirico, infatti, è di per sé *eteronomo*, ossia impone una sua legge, contrastante con quella della libertà. In sostanza, «L'autonomia della volontà è l'unico principio di ogni legge morale, e dei doveri a questa legge conformi: ogni *eteronomia* dell'arbitrio,

* Facoltà di Filosofia, Università Vita-Salute San Raffaele.

¹ I. Kant, *Critica della ragion pratica*, (a cura di) V. Mathieu, Milano, Rusconi, 1982 (ed. originale 1788), p. 215.

per contro, non solo non fonda alcuna obbligatorietà, ma, anzi, è contraria al suo principio e alla moralità del volere. [...] la legge morale non esprime null'altro che l'autonomia della ragion pura pratica, cioè della libertà»².

Un'impostazione analoga, di marca essenzialmente razionalista, si può riscontrare anche nella principale tradizione di studio empirico della morale che percorre il XX secolo; gli studi sullo sviluppo morale dei bambini, iniziati da Jean Piaget e proseguiti da Lawrence Kohlberg e dalla sua scuola, concepiscono lo sviluppo del giudizio morale in maniera analoga a quello dei giudizi cognitivi; quanto più il fanciullo acquisisce capacità cognitive complesse tanto più affina anche le sue capacità di giudizio morale, passando attraverso stadi codificati che lo conducono da una morale *convenzionale*, di carattere sostanzialmente eteronomo, a una morale *post-convenzionale*, che coincide di fatto con l'idea kantiana di un giudizio autonomo, imparziale e di portata universale. L'individuo moralmente maturo giudica da sé, applicando imparzialmente regole e principi, senza farsi dominare da sentimenti ed emozioni; il punto d'arrivo è perciò costituito una forma di universalismo razionalistico di tipo kantiano³.

Esiste tuttavia anche una linea minoritaria, ma saldamente attestata nella storia della filosofia, che invece concepisce la morale come una questione eminentemente sentimentale. La tradizione britannica del sentimentalismo, iniziata da Shaftesbury e sviluppata da autori come Hutcheson e Hume, ha posto l'accento sul carattere pratico della morale, ossia sul fatto che i giudizi morali siano in grado di motivare efficacemente l'azione, e avanzato dubbi sull'effettiva capacità della ragione di governare le nostre passioni. Ha perciò difeso il carattere "passionale" dei giudizi morali, ossia il loro essere frutto di un "senso morale", di una capacità di provare specifici sentimenti di tipo morale, come l'orgoglio e la benevolenza. Hume, in particolare, ha sostenuto che, poiché la ragione si occupa solo di questioni fattuali o logiche (ossia di relazioni tra idee), non è in grado di muovere all'azione e solo le passioni, ossia i desideri e le inclinazioni, possono determinarci ad agire: come egli sostiene in un passo famoso, «La ragione è, e deve solo essere, schiava delle passioni e non può rivendicare in nessun caso una funzione diversa da quella di servire

² Ivi, p. 217.

³ J. Piaget, *Il giudizio morale nel fanciullo*, Giunti Barbera, Firenze, 1972 (ed. originale 1932); L. Kohlberg, *The Philosophy of Moral Development. Moral Stages and the Idea of Justice*, Harper & Row, New York, 1981.

e obbedire a esse»⁴. Le regole morali non possono essere il prodotto della ragione, proprio perché sono in grado di produrre o di impedire le azioni; pertanto, è solo grazie alle impressioni o ai sentimenti suscitati in noi dalle qualità morali che possiamo stabilire differenze morali: «La morale, perciò, è più propriamente oggetto di sentimento che di giudizio»⁵.

Questa tradizione sentimentalista è stata largamente riscoperta dalla psicologia e dalle neuroscienze contemporanee, attraverso numerose ricerche che hanno sottolineato la centralità della dimensione emotiva e sentimentale nella generazione dei giudizi morali.

2. Neosentimentalismo

Nel corso degli ultimi decenni, psicologia morale, scienze cognitive e neuroscienze hanno dedicato una notevole attenzione al fenomeno della morale⁶; quest'attenzione si è concretizzata nella critica al paradigma razionalista e nella sua sostituzione con un modello alternativo che fa della morale un fenomeno essenzialmente affettivo. Particolare risonanza ha avuto la ricerca dello psicologo americano Jonathan Haidt che, sulla base di numerosi studi sulla psicologia delle valutazioni morali, ha proposto una teoria "socio-intuizionista" del giudizio morale. Secondo questo approccio, la morale si baserebbe essenzialmente su "intuizioni"; con questo termine, però, non si deve intendere la conoscenza razionale, di tipo non inferenziale, postulata da numerosi autori della tradizione filosofica, sia per i primi principi logici sia per quelli morali, ma piuttosto l'attivazione di una risposta emotiva di carattere automatico⁷. Un giudizio morale, in altri termini, sarebbe collegato all'attivarsi spontaneo, non deliberato, di un senso di approvazione o disapprovazione che si impone in quanto carico di un "peso" emotivo. L'attivazione di queste emozioni

⁴ D. Hume, *Trattato sulla natura umana*, (a cura di) E. Lecaldano, Laterza, Roma-Bari, 2002, (ed. originale 1740), p. 436.

⁵ Ivi, p. 497.

⁶ T. Nadelhoffer T., E. Nahmias, S. Nichols, (a cura di), *Moral Psychology: Historical and Contemporary Readings*, Wiley-Blackwell, Malden, 2010; S. Songhorian, *Etica e scienze cognitive*, Carocci, Roma, 2020.

⁷ W. Sinnott-Armstrong, *Moral Intuitionism Meets Empirical Psychology*, in T. Horgan, M. Timmons (a cura di), *Metaethics after Moore*, Oxford University Press, Oxford, 2006, pp. 339-365; A. Kauppinen, *Moral Intuition in Philosophy and Psychology*, in J. Clausen, N. Levy (a cura di), *Handbook of Neuroethics*, Springer, Dordrecht, 2015, pp. 169-183.

che generano i nostri giudizi avrebbe una base innata perché le emozioni fondamentali deriverebbero da meccanismi che sono stati selezionati evolutivamente, ma verrebbe poi plasmata e orientata dall'inserimento in una certa struttura sociale⁸; quest'ultima determinerebbe con maggiore precisione gli oggetti verso cui le emozioni positive e negative vengono indirizzate e soprattutto determinerebbe il prevalere dell'uno o dell'altro schema generale di risposta morale. Secondo la concezione di Haidt e colleghi⁹, infatti, vi sarebbero tre principali schemi di interpretazione morale che attivano diversamente i "recettori morali" presenti nel nostro cervello, scatenando reazioni emotive e giudizi morali. I tre modelli fondamentali sarebbero i seguenti: la *morale della comunità*, che insiste sulla priorità del gruppo e sulla lealtà dell'individuo nei suoi confronti; la *morale della divinità* che si caratterizza per l'insistenza su norme di purezza relative a cibo, mestruazioni, comportamenti sessuali e trattamento di cadaveri; la *morale dell'autonomia* che sottolinea l'importanza dei diritti individuali e la centralità della distinzione tra regole convenzionali e morali¹⁰. Questa distinzione, in realtà, sarebbe un fatto puramente culturale, caratteristico della società individualista; infatti, lo studio delle società non occidentali testimonia l'esistenza di regole che, pur appearing ai nostri occhi come delle mere convenzioni, svolgono un effettivo ruolo morale all'interno di quei contesti¹¹. Alle tre forme di pensiero morale corrispondono tre emozioni morali di base: la violazione di regole comunitarie, in particolare di regole di autorità, genera l'emozione del disprezzo (*contempt*), la violazione dei diritti individuali dà luogo soprattutto alla rabbia (*anger*) e la violazione di norme di purezza e sacralità genera il disgusto (*disgust*). Queste tre forme principali di pensiero morale, e le emozioni fondamentali che vi si collegano, sarebbero variamente presenti in tutte le morali umane, ma la prevalenza dell'una o dell'altra viene appunto determinata dall'influsso delle diverse culture; in sostanza, la morale sarebbe un costruito culturale,

⁸ J. Haidt J., *Nativism and Moral Psychology: Three Models of the Innate Structure That Shapes the Contents of Moral Norms*, in W. Sinnott-Armstrong (a cura di), *Moral Psychology. Volume 1: The Evolution of Morality: Adaptation and Innateness*, The MIT Press, Cambridge, Mass., 2008, pp. 319-343.

⁹ P. Rozin, L. Lowery, S. Imada, J. Haidt, *The CAD Triad Hypothesis: A Mapping Between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity)*, in «Journal of Personality and Social Psychology», 76, 1999, pp. 147-163.

¹⁰ E. Turiel, *The Development of Social Knowledge: Morality and Convention*, Cambridge University Press, Cambridge, 1983.

¹¹ J. Haidt, S.H. Koller, M.G. Dias, *Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?*, in «Journal of Personality and Social Psychology», 65, 1993, pp. 613-628.

a partire però da una serie di materiali condivisi che l'evoluzione ha messo a disposizione della società umana.

Le evidenze sperimentali cui si appoggia questa teoria sono numerose. Che il dominio della morale sia legato a intuizioni ed emozioni, anziché a ragionamenti, è stato suggerito inizialmente dagli studi psicologici che prevedono l'interferenza di un pesante carico cognitivo (come tenere a mente un numero molto lungo) durante l'esecuzione del compito di pronunciare giudizi morali; in questi studi non si verificano differenze nella qualità dei giudizi morali, rispetto ai soggetti di controllo, mentre tali differenze dovrebbero presentarsi, se i giudizi fossero l'esito di un ragionamento esplicito. Che la morale non si limiti all'ambito del danno arrecato ad altri¹², ma investa tutti gli ambiti nei quali vengono generate forti emozioni di disapprovazione, si può mostrare analizzando i casi che comportano innocue violazioni di tabù, ossia guardando ai modi in cui reagiamo ad azioni sconvenienti che non danneggiano nessuno. Negli esperimenti di Haidt, i soggetti erano posti di fronte a racconti che coinvolgono pratiche particolarmente cariche di significato emotivo, ad esempio pratiche sessuali trasgressive come l'incesto o l'antropofagia, e dovevano esprimere un giudizio in merito. I risultati mostrano che i soggetti tendono a fornire molto rapidamente dei giudizi morali negativi, ma sono per lo più incapaci di fornire giustificazioni per le loro conclusioni. Da un lato, le loro reazioni affettive consentono di predire il giudizio in maniera più affidabile rispetto alla percezione del carattere effettivamente dannoso delle pratiche, dall'altro l'incapacità di trovare ragioni plausibili per il giudizio non porta questi soggetti a modificarlo¹³. Haidt denomina questo fenomeno "ammutolire morale" (*moral dumbfounding*); la confusione mentale generata nei soggetti sperimentali da esempi di questo tipo è un'evidenza rilevante in favore della tesi per cui il giudizio morale non si basa sulla ponderazione di ragioni o l'applicazione di regole, ma sull'espressione di risposte spontanee, rapide e automatiche a fronte di situazioni moralmente problematiche. Una conferma dell'approccio intuizionista sociale verrebbe dalla verifica di una relazione causale tra esperienze visive e olfattive molto forti, come quelle di disgusto, e l'incremento della severità dei giudizi morali. Posti sotto l'influsso di un odore sgradevole, o messi a sedere a un tavolo molto sporco, i soggetti sperimentali forniscono giudizi molto più negativi su certi comportamenti; e quando vengono suggestionati a provare disgusto

¹² Come pensano altri studiosi: vedi E. Turiel, *The Development of Social Knowledge: Morality and Convention*, cit.

¹³ J. Haidt, M.A. Hersh, *Sexual Morality: The Cultures and Emotions of Conservatives and Liberals*, in «Journal of Applied Social Psychology», 31, 2001, pp. 191-221.

nell'udire una parola moralmente neutra, come "spesso" o "portare", i loro giudizi su azioni moralmente sbagliate peggiorano sensibilmente e possono arrivare a biasimare dei comportamenti che non sono in alcun senso sbagliati¹⁴.

Questa sorta di "neosentimentalismo" basato sulle ricerche empiriche distingue in maniera netta tra processi intuitivo-emotivi e processi razionali, affermando che i giudizi morali sono determinati dai primi. I processi riflessivi svolgono una funzione secondaria, di supporto al giudizio formulato in maniera indipendente sulla base delle emozioni; pertanto, assumono per lo più l'aspetto di una mera razionalizzazione, quando non di una vera e propria confabulazione, ossia del tentativo di rintracciare considerazioni, eventualmente incoerenti con la situazione, che potrebbero svolgere un ruolo giustificativo del nostro giudizio. Come scrive Haidt, «Nel modello intuizionista sociale, uno sente un rapido lampo di ripugnanza al pensiero dell'incesto e sa intuitivamente che c'è qualcosa di sbagliato. Poi, di fronte alla richiesta sociale di una giustificazione verbale, diventa un avvocato che cerca di costruire un argomento, anziché un giudice che cerca la verità»¹⁵. I processi razionali, in altri termini, non svolgono alcuna funzione *ex ante*, ossia non contribuiscono a determinare quale sarà il giudizio, o che cosa sia giusto pensare, ma solo una funzione *ex post*, ossia di sostegno giustificativo; un sostegno, peraltro, che non mira a convincere con argomenti ma a generare emozioni analoghe nell'interlocutore, onde persuaderlo della bontà della nostra conclusione: perciò, «Il nostro pensiero morale è molto più simile a quello di un politico a caccia di voti che a quello di uno scienziato alla ricerca della verità»¹⁶. Per dirlo con la metafora utilizzata da Haidt nel titolo del suo contributo più noto, i processi razionali non sono altro che l'agitarsi della coda del cane emotivo; un segno esteriore – vale a dire – che corrisponde a un evento interiore di rilievo sostanziale, evento che ha carattere emotivo.

Non solo, pertanto, i giudizi morali sono determinati pressoché univocamente dalle nostre reazioni emotive, ma sono anche assai poco inclini ad assumere una portata universale, contrariamente a quanto sostiene

¹⁴ T. Wheatley, J. Haidt, *Hypnotic Disgust Makes Moral Judgments More Severe*, in «Psychological Science», 16, 2005, pp. 780-784; S. Schnall, J. Haidt, G. L. Clore, A. H. Jordan, *Disgust as Embodied Moral Judgment*, in «Personality and Social Psychology Bulletin», 34, 2008, pp. 1096-1109.

¹⁵ J. Haidt, *The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, in «Psychological Review», 108, 2001, pp. 814-834, cit. p. 814.

¹⁶ J. Haidt, *Menti tribali. Perché le brave persone si dividono su politica e religione*, Codice Edizioni, Torino, 2013, cit. p. 99.

la tradizione morale occidentale. Da un punto di vista evoluzionistico, infatti, la morale rappresenta un mezzo per unire le persone in un gruppo, facendole partecipare a imprese collettive, tenendo a bada le loro inclinazioni egoistiche e scoraggiando i profittatori. In sostanza, è un potente fattore di coesione che rende possibili le società cooperative regolando l'interesse personale e consentendo il raggiungimento di un certo livello di benessere. L'altra faccia della medaglia, però, è l'intrinseca tendenza delle nostre disposizioni morali al "gruppismo", ossia a rinforzare i legami con i membri del nostro gruppo attraverso la contrapposizione dei nostri bisogni e interessi con quelli degli individui al di fuori del gruppo. La nostra psicologia limita significativamente la capacità di avere relazioni empatiche con persone molto distanti e molto diverse da noi, ci confina a un raggio spazio-temporale di azione relativamente breve e rende inutili gli appelli a una considerazione universalistica dei bisogni umani che si estenda a chi è lontano da noi nel tempo e nello spazio¹⁷.

Complessivamente, questa teoria tende a ridurre al minimo, se non ad annullare, lo spazio della ragione nella formulazione dei giudizi morali. In base all'immagine che essa ne fornisce, non è praticamente mai vero che deliberiamo *per arrivare* a una conclusione morale, soppesando i pro e i contro. Questa conclusione si scontra con l'esperienza morale ordinaria in cui capita per l'appunto di dover soppesare considerazioni in conflitto e di decidere sulla base di un calcolo, certamente fallibile ma non integralmente dettato dalle emozioni, di costi e benefici. La teoria neosentimentalista può sembrare convincente, da un lato, perché è formulata a partire da situazioni sperimentali in cui occorre prendere una decisione rapidamente, senza avere la possibilità di una riflessione distesa; in questi casi è del tutto plausibile che le euristiche mentali, i meccanismi automatici e le reazioni emotive giochino un ruolo primario o dominante, ma ciò non dimostra che lo stesso avvenga nei casi in cui la scelta non è sottoposta ai medesimi vincoli. D'altro canto, la tesi neosentimentalista trova conferme sperimentali perché tende per lo più a mettere in gioco situazioni estreme o anche poco consuete che generano forti emozioni e generalmente non rappresentano dei veri dilemmi: nelle opzioni che vengono fornite ai soggetti sperimentali c'è spesso un'opzione che appare decisamente preferibile, il che invita a seguire le proprie intuizioni iniziali, senza imbarcarsi in un più lungo ed esigente percorso di rivalutazione critica.

¹⁷ Ivi.

3. *La teoria del doppio processo*

Un'immagine leggermente più articolata del rapporto tra emozioni e processi riflessivi è stata offerta da un altro autore che ha esercitato un'influenza profonda sugli studi recenti: si tratta di Joshua Greene il quale ha applicato al giudizio morale la teoria del doppio processo elaborata da Daniel Kahnemann¹⁸. Questa teoria prevede un duplice livello di elaborazione cognitiva: i processi veloci, che attivano aree cerebrali prevalentemente di tipo emotivo e consentono di fornire risposte rapide e standardizzate, spesso basate su euristiche, pregiudizi e semplificazioni; i processi lenti che richiedono invece l'intervento di aree deputate alla computazione e alla riflessione ma consentono una maggiore sensibilità alle circostanze e risposte informate da un più preciso calcolo di costi e benefici. In un famoso articolo pubblicato nel 2001, Greene ha sottoposto ai soggetti sperimentali sotto risonanza magnetica funzionale sessanta dilemmi divisi in tre categorie: i dilemmi di carattere non morale, come scegliere se prendere l'autobus o il treno, considerati certi vincoli di tempo; i dilemmi morali di carattere "non personale", ossia quelli con uno scarso contenuto emotivo, e i dilemmi morali di carattere "personale", ossia quelli che comportano un rapporto diretto o anche un contatto fisico con un'altra persona e l'inflizione diretta di un danno¹⁹. Questi ultimi sono rappresentati paradigmaticamente dall'ipotesi – partorita dalla fervida fantasia di Judith Thomson – di spingere un uomo grasso giù da un ponte per bloccare un carrello in corsa che altrimenti ucciderebbe cinque persone; i dilemmi morali impersonali, invece, sono rappresentati dal gesto di tirare una leva di scambio per deviare il medesimo carrello dal binario in cui, proseguendo, ucciderebbe cinque persone a quello in cui ne ucciderebbe una sola. Se richiesta di operare una scelta, la maggior parte delle persone accetta di tirare la leva nel secondo caso, ma rifiuta di spingere l'uomo grasso nel primo; l'ipotesi interpretativa offerta da Greene e colleghi è che tale discrepanza dipenda dalla differenza tra dilemmi personali e dilemmi impersonali: i primi comportano l'esecuzione di un'azione che a) ha una forte probabilità di causare un danno fisico grave; b) riguarda una persona individuabile; c) non deriva dalla deviazione di una minaccia già

¹⁸ D. Kahneman, *A Perspective on Judgment and Choice: Mapping Bounded Rationality*, in «American Psychologist», 58, 2003, pp. 697-720; D. Kahneman, *Pensieri lenti e veloci*, Mondadori, Milano, 2020.

¹⁹ J. D. Greene, R. B. Sommerville, L. E. Nystrom, J.M. Darley, J.D. Cohen, *An fMRI Investigation of Emotional Engagement in Moral Judgment*, in «Science», 293, 2001, pp. 2105-2108.

esistente. In altri termini, la violazione morale qui immaginata presenta la struttura “io danneggio te” (*me hurt you*); la forte attivazione delle aree emotive causata da queste caratteristiche determina una risposta negativa alla domanda sulla liceità dell’atto, risposta che ha carattere automatico ed è molto difficile da sovrastare. Nei dilemmi impersonali, invece, non c’è questa forte attivazione emotiva e perciò il processo computazionale, affidato alle aree più “cognitive”, può più facilmente soverchiare una risposta emotiva più debole.

Nello studio di Greene venivano evidenziate attivazioni analoghe, relative ad aree di tipo cognitivo superiore, tra i dilemmi non morali e quelli morali non personali; nei dilemmi morali personali, invece, c’era una prevalente attività nelle aree deputate alle emozioni e alla cognizione sociale che tendeva a inibire la risposta in termini consequenzialisti. Il sistema emotivo, infatti, ha carattere automatico, fa ricorso a un’elaborazione parallela delle informazioni ed è perciò rapido nell’attivarsi e nel raggiungere conclusioni, mentre il sistema razionale, che elabora le informazioni in maniera seriale, richiede più tempo. Infatti, nei soggetti che forniscono risposte improntate a criteri consequenzialisti, si osserva un incremento consistente del tempo di reazione; incremento che viene interpretato come il risultato di una “battaglia” che la parte “razionale” deve intraprendere contro quella “emotiva” prima di poter imporre il proprio giudizio. Chi giudica appropriata la violazione personale presenta infatti un incremento di attività della corteccia dorsolaterale prefrontale, che svolge compiti di ragionamento astratto e controllo cognitivo; più in generale, chi ha un più lungo tempo di reazione nei dilemmi personali presenta un aumento sia dell’attività della corteccia cingolata anteriore, un’area nota per essere coinvolta nella gestione dei conflitti cognitivi, sia della corteccia dorsolaterale prefrontale. Tali incrementi sono maggiori nei dilemmi personali più difficili, che richiedono un tempo più lungo per processare l’evidenza.

Naturalmente l’osservazione di una correlazione tra certe risposte in certi soggetti e un consistente livello di attivazione delle aree emotive non dimostra di per sé l’esistenza di un nesso causale tra emozioni e giudizi; è possibile che la co-occorrenza sia spiegabile in altri modi. Tuttavia, varie evidenze che suggeriscono che l’interpretazione causale sia la migliore. Ad esempio, gli studi compiuti sui pazienti con lesione focale alla porzione ventromediale della corteccia prefrontale mostrano che questi pazienti – le cui emozioni sociali sono pressoché assenti – forniscono risposte utilitariste

in misura cinque volte maggiore rispetto ai controlli sani²⁰; inoltre, il tasso di risposte utilitariste è triplicato nei pazienti affetti da demenza frontotemporale, una condizione neurologica che è a sua volta associata a un appiattimento generale delle emozioni²¹; risultati analoghi sono stati poi ottenuti anche in soggetti psicopatici a basso tasso di ansietà²²; infine, elevate attivazioni dell'amigdala, anche indotte dall'assunzione di un farmaco, sembrano determinare una riduzione delle risposte consequenzialiste²³. In sostanza, sembra esservi una consistente mole di dati che indica che, nei pazienti che provano scarsa empatia, compassione e senso di colpa, pur mantenendo capacità sostanzialmente normali per quanto concerne l'intelligenza generale, il ragionamento e perfino la conoscenza verbale delle norme morali, i meccanismi di risposta razionale e calcolatrice non sono inibiti dai modelli automatici di risposta emotiva; perciò, questi pazienti trattano i contesti emotivamente salienti, di tipo "personale", in maniera analoga a quelli di tipo non morale, tendendo sempre a massimizzare il risultato positivo. Tutto ciò sembra giustificare la conclusione per cui, nei casi normali, la risposta non è meramente *accompagnata* da certe emozioni, ma è invece *causata* dai meccanismi di risposta automatica mediati dalle aree emotive.

Il rilievo prioritario della dimensione "personale" del danno è stato successivamente confermato da altri esperimenti su soggetti sani, che hanno evidenziato un incremento significativo delle risposte positive alla domanda se uccidere l'uomo grasso quando la situazione veniva modificata in maniera da attenuare quella dimensione: infatti, i soggetti che giudicavano accettabile spingere l'uomo grasso con le mani o usando un bastone erano, rispettivamente, il 3 e il 33% del totale, ma salivano al

²⁰ M. Koenigs, L. Young, R. Adolphs, D. Tranel, F. Cushman, M. Hauser, A. Damasio, *Damage to the Prefrontal Cortex Increases Utilitarian Moral Judgment*, in «Nature» 446, 2007, pp. 908-911; E. Ciaramelli, M. Muccioli, E. Làdavas, G. di Pellegrino, *Selective Deficit in Personal Moral Judgment Following Damage to Ventromedial Prefrontal Cortex*, in «Social Cognitive and Affective Neuroscience», 2, 2007, pp. 84-92.

²¹ M.F. Mendez, E. Anderson, J. S. Shapira, *An Investigation of Moral Judgment in Frontotemporal Dementia*, in «Cognitive and Behavioural Neurology», 18, 2005, pp. 193-197.

²² M. Koenigs, M. Kruepke, J. Zeier, J. P. Newman, *Utilitarian Moral Judgment in Psychopathy*, in «Social Cognitive and Affective Neuroscience», 7, 2012, pp. 708-714.

²³ A. S. Shenhav, J. D. Greene, *Integrative Moral Judgment: Dissociating the Roles of the Amygdala and the Ventromedial Prefrontal Cortex*, in «Journal of Neuroscience», 34, 2014, pp. 4741-4749; M-J. Crockett, *Serotonin Selectively Influences Moral Judgment and Behavior through Effects on Harm Aversion*, in «Proceedings of the National Academy of Sciences», 107, 2010, pp. 17433-17438.

59 e al 63% se la vittima veniva fatta cadere azionando una botola con un bottone, vuoi stando sul ponte vuoi stando altrove, lontano dalla vittima²⁴.

L'aver individuato un ruolo attivo nel giudizio morale sia per i meccanismi emotivi sia per quelli razionali distingue la posizione di Greene da quella di Haidt: mentre quest'ultimo pensa che nei giudizi morali si tratti sempre essenzialmente di intuizioni emotivamente cariche, Greene è disposto a concedere di più al pensiero riflessivo, ossia individua una serie di situazioni in cui i processi razionali hanno la meglio sulle risposte automatiche, generando conclusioni di tipo consequenzialista. Greene interpreta perciò deontologia e consequenzialismo non come due teorie filosofico-normative, ma come due generi naturali psicologici. Il primo si può intendere come una propensione a giudicare in base alle euristiche messe a disposizione dal sistema emotivo ed è supportato da processi cerebrali inconsapevoli e automatici; non costituisce perciò una forma di ragionamento morale vero e proprio, ma una mera razionalizzazione morale. Il secondo, invece, si basa su processi psicologici di tipo cognitivo, consapevoli e riflessivi e perciò è una forma di ragionamento morale genuino. In questo modo, «la tensione centrale in etica tra deontologismo e consequenzialismo è una manifestazione della tensione centrale nel progetto cognitivo tra efficienza e flessibilità»²⁵. I processi automatici sono efficienti perché si attivano rapidamente e non richiedono sforzo; essi risultano anche pienamente affidabili quando sono stati formati dall'esperienza (genetica, culturale o personale) per tentativi ed errori. Quando però ci troviamo di fronte a problemi morali inediti, come quelli che traggono origine da sviluppi tecnologici recenti, o a problemi in cui vi è una consistente area di disaccordo, secondo Greene dovremmo appoggiarci ai processi riflessivi e al calcolo di costi e benefici; infatti, pensare di poter trattare quei problemi con meccanismi automatici progettati per tutt'altri scopi significa sperare in una sorta di miracolo cognitivo. Con un'utile analogia, Greene paragona i meccanismi di risposta automatica, di tipo deontologico, all'utilizzo di una macchina fotografica in modalità automatica; le impostazioni preselezionate vanno bene in situazioni ordinarie, ma offrono soluzioni stereotipate, prive della necessaria flessibilità in casi nuovi e più complessi. Quando c'è bisogno di maggior flessibilità e raffinatezza di analisi, occorre utilizzare la modalità manuale: «Le nostre emozioni morali – le nostre

²⁴ J. D. Greene, F. A. Cushman, L. E. Stewart, K. Lowenberg, L. E. Nystrom, J. D. Cohen, *Pushing Moral Buttons: The Interaction Between Personal Force and Intention in Moral Judgment*, in «Cognition», 111, 2009, pp. 364-371.

²⁵ J.D. Greene, *Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics*, in «Ethics», 124, 2014, pp. 69-726, cit. 699.

impostazioni automatiche – sono generalmente appropriate per bloccare il semplice egoismo ed evitare la tragedia dei beni comuni. È ciò per cui sono state programmate, sia biologicamente sia culturalmente. Perciò, quando il problema è Io contro Noi (o Io contro Te), dovremmo affidarci alle nostre reazioni viscerali, altrimenti dette coscienza [...] Ma... quando si tratta della tragedia della morale di senso comune – quando è Noi contro Loro – è tempo di smetterla di prestare fiducia ai sentimenti viscerali e passare alla modalità manuale»²⁶.

Il modello del doppio processo prospetta pertanto una sorta di co-occorrenza di processi di tipo emotivo e di tipo riflessivo che agiscono in maniera indipendente e si trovano spesso in conflitto tra loro. Come osserva Greene, non è qui in questione propriamente un conflitto tra sentimenti e ragione, poiché in entrambi i casi si tratta di processare informazioni e quindi di processi che, in senso lato, possono definirsi cognitivi; si può parlare, invece, di una competizione tra sistema cognitivo *automatico* e sistema cognitivo *deliberato*, ovvero di «un contrasto tra rappresentazioni che hanno una forza motivazionale diretta, da un lato, e dall'altro rappresentazioni che non hanno una forza motivazionale diretta loro propria, ma che possono essere contingentemente connesse agli stati affettivi/emotivi che possiedono una tale forza, producendo così un comportamento flessibile e diretto allo scopo»²⁷. Questa concezione più equilibrata dei rapporti tra processi emotivi e processi riflessivi è stata in parte confermata da ulteriori esperimenti che hanno replicato l'analisi delle risposte e dei tempi di reazione confrontando due gruppi, uno dei quali doveva svolgere un compito cognitivo supplementare, consistente nello schiacciare un bottone ogniqualvolta compariva il numero 5 su uno schermo in cui venivano proiettati 3,5 numeri al secondo; la presenza di questo carico cognitivo determinava un significativo incremento dei tempi di reazione in chi dava risposte utilitariste, rispetto al gruppo di controllo, mentre non vi era alcun incremento del genere nei soggetti sottoposti al carico cognitivo che davano risposte non utilitariste²⁸. Queste differenze

²⁶ J. D. Greene, *Moral Tribes: Emotion, Reason and the Gap Between Us and Them*, Penguin, New York, 2013, cit. p. 294.

²⁷ J. D. Greene, L. E. Nystrom, A. D. Engell, J. M. Darley, J. D. Cohen, *The Neural Basis of Cognitive Conflict and Control in Moral Judgment*, in «Neuron», 44, 2004, pp. 389-400, cit. p. 397-398.

²⁸ J. D. Greene, S. A. Morelli, K. Lowenberg, L. E. Nystrom, J. D. Cohen, *Cognitive Load Selectively Interferes with Utilitarian Moral Judgment*, in «Cognition», 7, 2008, pp. 1144-1154; J. D. Greene, F. A. Cushman, L. E. Stewart, K. Lowenberg, L. E. Nystrom, J. D. Cohen, *Pushing Moral Buttons: The Interaction Between Personal Force and Intention in*

si sono registrate sia presso chi, nel complesso, tendeva a dare risposte utilitariste, sia in chi propendeva per risposte non utilitariste; l'aggiunta del carico cognitivo, peraltro, non diminuiva la percentuale delle risposte utilitariste osservate, come pure si sarebbe potuto prevedere. Questi risultati sembrano costituire una conferma del nesso tra giudizi utilitaristi e processi cognitivi controllati da un lato e giudizi non utilitaristi e processi automatici dall'altro, e sembrano altresì confermare l'esistenza di un qualche ruolo dei processi controllati anche nel giudizio morale *ex ante*, e quindi non in funzione di mera razionalizzazione.

4. *Aspetti critici e approcci alternativi*

I due modelli che abbiamo richiamato hanno esercitato un'influenza notevole nella discussione recente. Entrambi sono tuttavia oggetto di possibili critiche perché sembrano recepire una concezione ancora troppo dicotomica del rapporto tra emozioni e processi riflessivi. Una simile dicotomia, come si è notato, è da sempre parte della tradizione filosofica occidentale: se per un verso il razionalismo ha teso a sottovalutare il contributo della dimensione emotiva e sentimentale, squalificandola come estranea a una dimensione genuinamente morale, per altro verso il sentimentalismo ha per lo più visto nei processi razionali il semplice rispecchiamento di fatti empirici, privo di ogni rilievo valoriale e incapace di generare motivazione. Tale dicotomia merita però di essere problematizzata; e la stessa riflessione empirica contemporanea offre ragioni per formulare una rappresentazione molto più collaborativa del rapporto tra i due elementi, respingendo da un lato la caratterizzazione dei processi emotivi come meramente passivi e insensibili alle ragioni, dall'altro quella dei processi riflessivi come «freddi» e calcolanti. Sebbene queste caratterizzazioni possano avere una limitata plausibilità all'interno di certi contesti, e in particolare nelle situazioni tipicamente oggetto della ricerca empirica, dove si devono prendere decisioni molto rapide e in situazioni decisamente non ordinarie, si può dubitare del loro valore nei contesti di giudizio morale che si presentano nell'esperienza ordinaria.

Il fatto che le reazioni emotive abbiano un ruolo rilevante nel giudizio

Moral Judgment, cit.; B. Trémolière, W. De Neys, J.- F. Bonnefon, *Mortality Salience and Morality: Thinking About Death Makes People Less Utilitarian*, in «Cognition», 124, 2012, pp. 379-384.

e nella presa di decisione morale non è in discussione; ciò che invece merita di essere approfondito è, da un lato, il carattere “originario” e non derivato di tali reazioni, dall’altro la loro reale contrapposizione ai processi di tipo riflessivo. Le emozioni morali e le nostre disposizioni ad approvare o disapprovare hanno un carattere, per così dire, “primo”? Sono – per dirla con Hume – delle “esistenze originarie”, ossia dei fatti completi in se stessi che non rappresentano qualcosa al di là di se stessi né implicano alcun riferimento ad altro da sé? È lecito dubitarne; in realtà, le nostre stesse emozioni spontanee sembrano essere a loro volta “cariche di teoria”, ovvero di giudizi e di processi cognitivi²⁹. Vi è cioè un ruolo dell’esperienza precedente, sedimentata in giudizi di carattere generali e di regole apprese attraverso l’educazione, nel determinare la forma assunta dalle nostre reazioni emotive; al tempo stesso, le nostre capacità riflessive consentono – soprattutto se disponiamo di tempo – di porre una distanza critica tra noi e le reazioni emotive che ci capita di manifestare, così da rivalutarne la qualità e modificarne in parte le forme. Esiste, in altri termini, una capacità di controllo cognitivo delle emozioni che è stata oggetto di attenti studi psicologici e neuroscientifici. In particolare, i ricercatori hanno individuato cinque modalità attraverso le quali tale controllo può essere effettuato³⁰: in primo luogo, attraverso la selezione delle situazioni che ci consente di evitare certi stimoli emotivi; in secondo luogo, mediante la modifica delle situazioni che consente di attutire o rimuovere il rischio che si generino determinate emozioni; in terzo luogo, grazie al dislocamento dell’attenzione che porta a focalizzarsi su stimoli emotivamente non “pericolosi”; in quarto luogo, mediante un cambiamento cognitivo, ossia una rivalutazione di elementi della situazione che consenta di concettualizzarla in maniera diversa; infine, con la modulazione della risposta che interviene a modificare l’esperienza emotiva in corso, attenuandone la portata. La manipolazione cognitiva delle emozioni attraverso questi processi si è dimostrata efficace in molte circostanze, in particolare nel ridurre – a parità di attivazione fisica – l’esperienza soggettiva di emozioni negative come il disgusto che, come si è visto in precedenza, giocano un ruolo rilevante anche nella formulazione di giudizi morali; e l’uso di queste strategie di controllo può divenire automatico grazie all’adozione di schemi mentali ricorrenti.

È importante osservare che queste strategie cognitive di rivalutazione

²⁹ D. A. Pizarro, P. Bloom, *The intelligence of the moral intuitions: A comment on Haidt (2001)*, in «Psychological Review», 110, 2003, pp. 193-196.

³⁰ K. N. Ochsner, J. J. Gross, *Cognitive emotion regulation: Insights from social cognitive and affective neuroscience*, in «Current Directions in Psychological Science», 17, 2008, pp. 153-158.

delle emozioni sembrano basarsi su relazioni bidirezionali tra aree cerebrali associate alle emozioni e aree associate a compiti cognitivi³¹. Il che, complessivamente, suggerisce che il rapporto tra stati affettivi e processi riflessivi sia più complesso di quanto non appaia dai modelli basati sul doppio processo e che il modo in cui le reazioni emotive che tipicamente li accompagnano influenzano i giudizi morali possa essere significativamente variabile in diverse circostanze. In luogo di una determinazione univoca dei giudizi da parte dei processi emotivi, determinazione eventualmente modificabile da un successivo intervento di processi computazionali, sembra delinearsi una concezione collaborativa nella quale elementi affettivi ed elementi cognitivi convergono nel determinare una regolazione emotiva che può andare sia nella direzione di mitigare l'eccitazione sia in quella di alimentarla. Le differenze osservate da Greene tra ciò che accade nei dilemmi morali personali e in quelli impersonali potrebbero forse non derivare dal contrasto tra due meccanismi indipendenti e in competizione tra loro, ma essere invece l'effetto dell'attivazione di due diverse strategie regolative delle emozioni: il caso del dilemma personale tende ad attivare strategie di controllo che mirano ad alimentare la reazione emotiva, quello impersonale induce strategie tese a mitigarla³². Il modo in cui veniamo effettivamente influenzati dalle emozioni, in altri termini, dipende anche da fattori cognitivi preesistenti; è perciò che gli individui che hanno una preesistente inclinazione verso il consequenzialismo tendono a provare minori emozioni anche nel caso personale, dal momento che sono già predisposti ad attivare strategie di controllo che le riducono³³.

Queste evidenze tratte dalle neuroscienze delle emozioni hanno condotto vari autori a proporre modelli di interpretazione del giudizio morale che rivalutano in misura consistente il ruolo svolto dai processi intuitivi, senza metterlo in contrapposizione con quello dei processi riflessivi. Peter Railton, ad esempio, ha suggerito che un modello alternativo per caratterizzare i processi intuitivi può essere quello di intenderli nei termini di una competenza implicita, analoga a quella tipicamente posseduta da un parlante nativo nei confronti della propria lingua madre; le intuizioni incorporano una serie di elementi percettivi, cognitivi ed

³¹ T. D. Wager, M. L. Davidson, B. L. Hughes, M. Lindquist, K.N. Ochsner, *Prefrontal subcortical pathways mediating successful emotion regulation*, in «Neuron», 59, 2008, pp. 1-14.

³² C. Helion, D. A. Pizarro, *Beyond Dual-Processes: The Interplay of Reason and Emotion in Moral Judgment*, in J. Clausen, N. Levy (a cura di), *Handbook of Neuroethics*, Springer, Dordrecht, 2015, pp. 109-125.

³³ D. M. Bartels, D. A. Pizarro, *The Mismeasure of Morals: Antisocial Personality Traits Predict Utilitarian Responses to Moral Dilemmas*, in «Cognition», 121, 2011, pp. 154-161.

emotivi che, senza dover essere oggetto di un'attenzione consapevole e deliberata, rendono possibile un'agevole navigazione nel mare complesso della vita ordinaria e delle relazioni sociali. Sulla scia dei risultati empirici più recenti, secondo Railton, è possibile caratterizzare il sistema affettivo come «un sistema per processare le informazioni flessibile e basato sull'esperienza, pienamente in grado di rintracciare correlazioni statistiche e di guidare le scelte comportamentale attraverso il bilanciamento di costi, benefici e rischi»³⁴. Le emozioni, in particolare, sono modelli di risposta selezionati evolutivamente per incrementare la capacità degli esseri umani di fronteggiare le sfide adattative; esse incorporano meccanismi di risposta che si sono rivelati utili per consentire agli esseri umani di sintonizzarsi sugli elementi del loro ambiente e rispondere ad esso in maniera efficace. Pertanto, esse possono essere caratterizzate come «risposte a stimoli interni ed esterni evolutesi per sintonizzare l'organismo in maniera sistemica rispetto alle sfide e alle opportunità che incontra nel corrispondere ai propri bisogni e nel realizzare i propri scopi in circostanze fisiche e sociali mutevoli»³⁵. Riprendendo la metafora di Haidt e giocando con l'omofonia dei termini *tail* (coda) e *tale* (storia), Railton ha pertanto suggerito che, a ben vedere, il cane affettivo racconti in verità una storia razionale; ovvero sia che le nostre reazioni emotive contengano un'ampia serie di competenze sociali e valoriali senza le quali non saremmo in grado di orientarci nel complesso panorama normativo delle nostre vite.

Un'analogia immagine revisionista del rapporto tra emozioni e processi riflessivi nella formulazione dei giudizi morali è stata proposta anche da Hanno Sauer che ha sottolineato come le emozioni siano educabili attraverso l'esperienza³⁶. Da un lato, infatti, attraverso la scelta degli input situazionali con cui confrontarsi, le nostre reazioni emotive si arricchiscono e vengono plasmate in nuove direzioni, dall'altro, grazie alla capacità metacognitiva di monitorare riflessivamente le nostre operazioni cognitive, è possibile determinare dei feed-back nei processi emotivi, modificandoli parzialmente. In altri termini, benché sia vero che i nostri giudizi morali sono basati su processi automatici e che il ragionamento morale consapevole viene spesso dopo il giudizio e la scelta morale, questo non significa che il razionalismo morale sia falso, nella misura in cui le

³⁴ P. Railton, *The Affective Dog and Its Rational Tale: Intuition and Attunement*, in «Ethics», 124, 2014, pp. 813-859, cit. p. 833.

³⁵ Ivi, p. 834.

³⁶ H. Sauer, *Educated intuitions. Automaticity and rationality in moral judgement*, in «Philosophical Explorations», 15, 2012, pp. 255-275; H. Sauer, *Moral Judgements as Educated Intuitions*, The MIT Press, Cambridge, Mass, 2017.

nostre reazioni automatiche incorporano processi di revisione critica e valutazione razionale delle nostre intuizioni. In altri termini, le intuizioni sono permeabili ai processi cognitivi, in particolare ai processi di revisione critica che hanno luogo nei contesti di interazione sociale; la capacità di ragionamento morale è lo sviluppo di un processo che progressivamente ci abitua a mettere in discussione e ad affinare le intuizioni che ci guidano nella vita ordinaria. Si tratta di «un processo costante esteso sia temporalmente sia intersoggettivamente, di formazione di giudizi intuitivi e di ragionamento motivante, che conduce a un miglioramento permanente, intersoggettivo e impegnativo delle proprie intuizioni, a un miglioramento del proprio ragionamento e a un miglioramento dell'interazione tra cognizione intuitiva e cognizione controllata»³⁷.

5. Conclusioni

Queste prospettive alternative delineano un'immagine del rapporto tra emozioni e processi riflessivi, ovvero tra moduli di risposta automatica e moduli di risposta deliberata, che, contrariamente a quella tradizionale, ha carattere non conflittuale: intuizioni ed emozioni non sono elementi passivi, meramente reattivi e irrazionali, ma sono in molti casi «cariche di riflessione» e perciò possono spesso costituire delle espressioni affidabili della nostra intelligenza pratica. Il fatto che si attivino spontaneamente, come dei meccanismi automatici, non toglie che incorporino dei processi razionali che non devono più passare attraverso la considerazione consapevole e riflessiva solo perché sono divenuti una sorta di «seconda natura». Come la riflessione etica sulle virtù, fin da Aristotele, aveva intuito, la ripetizione dei comportamenti e il consolidamento delle abitudini non indicano necessariamente o soltanto mancanza di pensiero e scarsa consapevolezza nell'agire, ma possono anche essere segnali della stabilizzazione di risposte appropriate e «sintonizzate». D'altro canto, i processi emotivi possono anche suggerire percorsi «lenti» di valutazione razionale, costituendo l'inizio di una riflessione volta a cogliere elementi delle situazioni che emergono soltanto da una prospettiva coinvolta e non distaccata.

La tradizione filosofica ha enfatizzato il ruolo dei processi riflessivi; molta parte della ricerca empirica contemporanea quello dei processi automatici. Entrambe, però, sembrano aver presupposto una fondamentale

³⁷ H. Sauer, *Moral Judgements as Educated Intuitions*, cit., p. 123.

dicotomia tra processi emotivi e processi riflessivi. Sulla scia di una parte consistente della riflessione più recente, si può ritenere che esistano buone ragioni per superare questa dicotomia e per ripensare le capacità morali a partire da un modello di tipo collaborativo e integrato tra processi «veloci» e processi «lenti», tra cognizioni affettivamente cariche e cognizioni più distaccate. In questa prospettiva, si può finalmente riconoscere che «il rapporto tra emozioni e cognizione non è tanto una lotta quanto una danza armoniosa tra partner paritetici e inseparabili; quando uno si sposta, l'altro si muove di conseguenza e sebbene uno possa assumere la guida, bisogna essere in due per ballare il tango»³⁸.

³⁸ C. Helion, K. N. Ochsner, *The role of emotion regulation in moral judgment*, in «Neuroethics», 11, 2018, 297- 308, cit. p. 298.

Diritto e neuroscienze: un avvicinamento a due velocità

FLAVIA CORSO*, ANDREA LAVAZZA**

SOMMARIO: 1. Premessa – 2. Neurodiritto, definizione e contesto – 2.1. La prospettiva neuroscientifica – 2.2. “Trattare” il crimine – 3. La realtà giuridica e del senso comune – 3.1. Il compatibilismo – 3.2. La tesi antinaturalistica – 3.3. Il ruolo delle emozioni – 4. Il situazionismo – 5. Conclusione.

1. *Premessa*

L'incontro del diritto con le neuroscienze è piuttosto recente. E soltanto all'alba di questo secolo si sono fatti espliciti tentativi di fare reagire le categorie giuridiche più rilevanti, soprattutto in ambito penale, con le scoperte più innovative sulla struttura e il funzionamento del nostro cervello. Il primo impatto tra le due discipline, promosso principalmente dal versante scientifico, non è avvenuto nel segno della collaborazione interdisciplinare ma di un tentativo quasi “imperialistico” delle neuroscienze di dettare una riforma ritenuta doverosa alla luce delle acquisizioni ottenute in laboratorio. Gli esseri umani analizzati attraverso il prisma del loro sistema nervoso sarebbero infatti molto diversi da quelli considerati dal diritto attraverso la lente della psicologia intenzionale e ciò imporrebbe di correggere gli istituti del diritto che non sono più adeguati alle conoscenze disponibili.

In questo senso la contrapposizione è posta tra una visione naturalistica, che mette al centro la descrizione scientifica delle basi materiali della nostra vita mentale, e una prospettiva ancorata alla concettualizzazione dei moventi umani in termini di credenze, desideri, decisioni considerati in astratto e attribuiti a un soggetto pienamente libero e razionale nell'orientare la propria condotta.

La prima, breve fase di entusiasmo per una potenziale rivoluzione alle porte ha lasciato il campo a un confronto più pacato e scientificamente equilibrato, in cui anche i giuristi e i filosofi del diritto e gli psicologi giuridici hanno portato il loro contributo e permesso alla discussione di virare verso

* University of Surrey.

** Centro universitario internazionale, Arezzo | Università di Pavia.

una più produttiva cautela rispetto a sovversioni improvvise di una lunga stratificazione nell'ambito della giustizia e della sua amministrazione.

Quello che abbiamo di fronte per il futuro prossimo sembra essere (e pensiamo che dovrebbe essere) una progressiva e sorvegliata apertura del diritto alle neuroscienze sia in ambito fondativo sia in ambito giurisprudenziale, senza però dimenticare che la scienza stessa è in evoluzione e le sue scoperte sono, per definizione, provvisorie e rivedibili. Ciò non significa che non vi sia progresso e tutte le nozioni su cui ci basiamo siano instabili. Una obiettiva lettura del cammino della scienza e del miglioramento medio della condizione umana ci dice il contrario. Singoli aspetti, tuttavia, soprattutto in un campo recente e in tumultuoso avanzamento come quello delle discipline del cervello, possono essere, e non di rado sono, corretti, precisati o del tutto superati.

Nel presente capitolo proponiamo una succinta introduzione del cosiddetto neurodiritto e mostriamo come vi siano attualmente diverse interpretazioni, allo stato tutte legittime, dell'approccio naturalistico-neuroscientifico. Inoltre, portiamo due esempi che approfondiscono la discussione di tale prospettiva. Il primo riguarda l'interpretazione dei dati neurali rispetto all'attribuzione di libertà al soggetto agente. Il secondo è un caso di revisionismo scientifico che indica come l'automatica applicazione dei risultati sperimentali può risultare fuorviante. Lo facciamo con il cosiddetto situazionismo, una teoria nell'alveo della psicologia empirica, per indicare che anche da un settore più consolidato rispetto alle neuroscienze sono da esportare e incorporare con prudenza acquisizioni e letture che promettono svolte nella secolare "saggezza" del diritto. E ciò non è dovuto a un conservatorismo epistemico o, peggio, sociale e culturale, ma a un dato di realismo e di consapevolezza scientifica.

2. *Neurodiritto, definizione e contesto*

Il termine "neurodiritto" (dall'inglese *neurolaw*) fa riferimento a quel campo di studi interdisciplinare che si pone l'obiettivo di studiare se e in che modo le neuroscienze possono influenzare le categorie concettuali su cui si fonda il diritto, in particolare il diritto penale¹.

Il punto di svolta per il rapporto tra neuroscienze e diritto si sviluppa a

¹ Per un'analisi dettagliata del neurodiritto, si vedano, tra gli altri, A. Lavazza e L. Sammiceli, *Il delitto del cervello. La mente tra scienza e diritto*, Torino, Codice Edizioni, 2012; e E. Picozza et al., *Neurodiritto. Una introduzione*, Torino, Giappichelli, 2014.

partire dalla fine del XX secolo, in continuità con le progressive innovazioni nell'ambito della scienza forense del secolo precedente², fino ad arrivare alle più recenti scoperte che concernono le basi fisiologiche e neuronali del comportamento deviante.

Tra i fattori che hanno permesso agli scienziati di approfondire lo studio inerente al rapporto tra cervello e comportamento (soprattutto criminale), un impatto notevole è riconducibile all'utilizzo di nuovi strumenti come l'elettroencefalogramma (EEG) e la risonanza magnetica funzionale (fMRI); queste tecnologie (in particolar modo la fMRI) consentono di individuare quali aree cerebrali si attivano (o non si attivano) in concomitanza con l'insorgere di uno specifico comportamento o manifestazione emotiva. In altri termini, le tecniche di *neuroimaging* permettono agli scienziati di superare la barriera della scatola cranica e indagare, seppure non ancora direttamente, ciò che avviene al suo interno nel momento esatto in cui si verifica la condotta specifica oggetto di studio.

Nel corso del tempo, questi studi hanno fornito conoscenze del tutto nuove in merito alle (con)cause biologiche sottostanti la condotta che rientra nelle fattispecie previste dalla legge. In particolare, si è potuto evidenziare come il fenomeno deviante sia spesso associato a specifiche varianti genetiche, a malformazioni o anomalie cerebrali, e a un alterato equilibrio bio-chimico.

Un esempio, tra i numerosi possibili nell'ambito degli studi genetici, è la scoperta di varianti a bassa attività del gene MAOA, che si ritengono associate a una maggiore predisposizione a sviluppare un comportamento violento³. Secondo alcuni studi⁴, anche una conformazione cerebrale particolare (si segnalano in particolare anomalie dell'amigdala, dell'insula e del cingolato

² A questo proposito, si veda ad esempio A. Petoft e M. Abbasi, *A Historical Overview of Law and Neuroscience: From the Emergence of Medico-Legal Discourses to Developed Neurolaw*, in «Journal of European History of Law», vol. 10, n. 2, 2019, pp. 15-33.

³ Si veda, ad esempio, G. Guo, et al., *The VNTR 2 Repeat in MAOA and Delinquent Behavior in Adolescence and Young Adulthood: Associations and MAOA Promoter Activity*, in «European Journal of Human Genetics», vol. 16, n. 5, 2008, pp. 626-634; D. A. Stetler, et al., *Association of Low-Activity MAOA Allelic Variants with Violent Crime in Incarcerated Offenders*, in «Journal of Psychiatric Research», vol. 58, 2014, pp. 69-75; R. Tikkanen, et al. *Psychopathy, PCL-R, and MAOA Genotype as Predictors of Violent Reconvictions*, in «Psychiatry Research», vol. 185, n. 3, 2011, pp. 382-386.

⁴ Cfr. M. Boccardi, et al., *Cortex and Amygdala Morphology in Psychopathy*, in «Psychiatry Research: Neuroimaging», vol. 193, n. 2, 2011, pp. 85-92; Y. Gao e A. Raine, *Successful and Unsuccessful Psychopaths: A Neurobiological Model*, in «Behavioral Sciences & the Law», vol. 28, n. 2, 2010, pp. 194-210; A. L. Glenn e A. Raine, *Neurocriminology: Implications for the Punishment, Prediction and Prevention of Criminal Behaviour*, in «Nature Reviews Neuroscience», vol. 15, n. 1, 2014, pp. 54-63.

anteriore) sembrerebbe influire negativamente sulla capacità dei soggetti che ne sono affetti di controllare gli impulsi aggressivi, aumentando il rischio di porre in atto modelli comportamentali antisociali⁵. Inoltre, un ruolo decisivo in tal senso lo svolgerebbero anche alcuni neurotrasmettitori; in particolare, è stato mostrato come una alterazione nella funzione di queste sostanze chimiche (ad esempio, un livello di serotonina più basso rispetto alla norma) possa contribuire a condizionare il comportamento in senso antisociale.

La raccolta di dati empirici sul (mal)funzionamento cerebrale e le sue implicazioni sul processo decisionale dell'individuo hanno potenziali ripercussioni sui capisaldi teorici della sanzione penale⁶, articolandosi peraltro in differenti direzioni.

Un primo punto focale riguarda l'interpretazione della reale portata dei dati neuroscientifici. A questo proposito, si possono individuare, in linea di massima, due approcci distinti: il filone di coloro che si potrebbero chiamare *moderati* e di coloro che si potrebbero chiamare *radicali*⁷. L'approccio dei moderati è caratterizzato dalla convinzione che, sebbene i risultati della scienza possano certamente fornire nuovi spunti di riflessione in ambito giuridico, la portata di questi risultati non sia tale da rivoluzionare l'impianto del diritto. Sul versante opposto, i radicali ritengono che le scoperte neuroscientifiche non siano solo potenzialmente in grado di riscrivere il fondamento del diritto penale, ma che debbano necessariamente farlo per motivi morali, oltre che per ragioni di aderenza del diritto con la realtà del funzionamento del mondo e degli esseri umani che ne sono parte integrante e non distinta⁸.

⁵ Non può sfuggire che termini come "anomalia", "anormalità", "alterazioni" o "scostamenti dalla media" siano centrali nella traduzione dal dato puramente quantitativo dell'analisi clinico-scientifica a una comprensione semantica contestualizzata. Ma essi costituiscono anche termini valutativi carichi di significati da spiegare e circoscrivere ulteriormente, pena una distorsione del quadro interpretativo. Infatti, dire "anormale" implica una "normalità" di riferimento che, nel caso citato del MAOA, sia anche il cardine dell'imputabilità del soggetto. Se per alcune patologie questa delimitazione è consolidata, per aspetti del tutto nuovi, come quelli genetici, fissare parametri è preliminare a possibili utilizzi giurisprudenziali di queste diagnosi. E fissare parametri appare un difficile compito da condividere tra esperti di genetica ed esperti di diritto in una complessa negoziazione tra diversi ambiti disciplinari e approcci scientifici.

⁶ A questo proposito si veda l'articolo di J. Greene e J. D. Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, in «Philosophical Transactions of the Royal Society B: Biological Sciences», vol. 359, n. 1451, 2004, pp. 1775-1785.

⁷ Cfr. A. Lavazza e L. Sammiceli, *Il delitto del cervello*, cit., p. 112.

⁸ Come si vedrà, ci si riferisce alla presunta impossibilità di giustificare la pena stante l'inesistenza della libertà umana e quindi della responsabilità personale per gli atti compiuti.

Vi è poi la questione relativa all'eventualità che, a prescindere da qualsivoglia lettura in merito alla portata delle neuroscienze, di fatto il diritto proceda (e debba procedere) secondo i propri criteri e concetti normativi, cosicché scienza e diritto costituirebbero due ambiti indipendenti gli uni dagli altri e non intersecabili tra loro se non in misura minore, soprattutto per quanto riguarda i fondamenti, mentre nel processo i punti di contatto potrebbero essere più estesi.

In ogni caso, se si prendono sul serio le recenti acquisizioni neuroscientifiche e alcune loro interpretazioni, ci si trova a dovere fare i conti con una nuova "realtà neuroscientifica" che sembra mettere in crisi, perlomeno *prima facie*, l'impalcatura teorica del diritto⁹. In questa sede, si intende perciò cominciare con l'analisi di questo punto di vista inedito, mettendo in luce quegli aspetti delle neuroscienze potenzialmente sovvertitori per l'ambito penale in primo luogo.

Si procederà dunque delineando dapprima i contorni della prospettiva neuroscientifica sull'essere umano, per poi evidenziare in che modo tale prospettiva si vada a scontrare, almeno in parte, con la visione implicita nel diritto e nel senso comune¹⁰.

2.1. *La prospettiva neuroscientifica*

Gli studi condotti nel campo delle neuroscienze e della genetica ci restituiscono un'immagine dell'essere umano per molti versi controintuitiva, che parte dall'equiparazione di mente e cervello su presupposti materialistici e continua mettendo in discussione libertà, razionalità e consapevolezza del nostro comportamento. Va oltre gli scopi di questo capitolo dare un quadro di tutto ciò. Ci limiteremo pertanto a considerare una serie di esperimenti sul libero arbitrio che hanno segnato il dibattito da quarant'anni a questa parte.

Gli studi condotti da Libet e colleghi¹¹ a partire dal 1983 – che costituiscono uno spartiacque in tal senso – hanno aperto la disorientante eventualità che le nostre convinzioni millenarie di senso comune sulla libertà umana possano essere in larga parte illusorie. I risultati di questi

⁹ Una questione epistemica non trascurabile riguarda l'interpretazione del dato neuroscientifico prima ancora della sua applicazione. Non è possibile qui analizzare il tema: lo si segnala però come rilevante.

¹⁰ Cfr. A. Lavazza e L. Sammiceli, *Il delitto del cervello*, cit., capp. 1-2.

¹¹ A questo proposito, si veda B. Libet, *Mind Time: The Temporal Factor in Consciousness*, Cambridge, Harvard University Press, 2004; trad. it. *Mind Time. Il fattore temporale nella coscienza*, Raffaello Cortina, Milano, 2007.

esperimenti sembrano infatti mettere in dubbio la concezione secondo cui gli esseri umani sono in grado di prendere le decisioni in maniera consapevole e razionale. Al contrario, negli studi di Libet si mostra che la consapevolezza dei soggetti di compiere una determinata azione (flettere un dito o piegare il polso) risulta posticipata di circa un terzo di secondo rispetto all'attivazione delle aree cerebrali preposte al cominciamento effettivo dell'azione stessa.

In sostanza, i dati empirici paiono indicare che le nostre scelte non sono determinate da una deliberazione cosciente del soggetto, bensì da meccanismi cerebrali inconsci su cui l'individuo non può esercitare alcun tipo di controllo. Ciò comporta necessariamente, secondo le interpretazioni più radicali di alcuni studiosi¹², una ridiscussione profonda del concetto di libero arbitrio, con il conseguente ridimensionamento dell'idea secondo cui l'essere umano è un agente morale dotato di autonomia e autodeterminazione. Tale presupposto, com'è noto, sottende tutte le nostre interazioni personali ed è alla base di gran parte delle pratiche sociali nonché degli istituti giuridici di quasi tutti gli ordinamenti.

Sulla base dell'*hard incompatibilism*¹³, che prevede che il determinismo sia reale e sia quindi incompatibile con il concetto di libero arbitrio, le implicazioni di una "convalida" neuroscientifica di questa controversa tesi sono tali da stravolgere completamente la struttura del diritto penale. E non solo: a mutare radicalmente sarebbe l'idea stessa della pena, la sua legittimità e la sua funzione¹⁴.

Il modello retributivo tradizionale – che permea anche il nostro sistema penale "misto" – trae infatti la propria giustificazione dal presupposto che gli individui, essendo agenti morali liberi, siano responsabili della propria condotta e, in quanto tali, possano essere biasimati e puniti. Tuttavia, se queste credenze non trovano effettiva corrispondenza nella realtà e il libero

¹² Tra le posizioni più radicali, si annoverano per esempio quelle di Greene e Cohen (J. Greene e J. D. Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, cit., e di Pereboom e Caruso (D. Pereboom e G. Caruso, *Hard-Incompatibilist Existentialism: Neuroscience, Punishment, and Meaning in Life*, in G. Caruso e O. Flanagan (eds.), *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, Oxford, Oxford University Press, 2017).

¹³ La tesi dell'*hard incompatibilism* è stata avanzata in D. Pereboom, *Determinism at Dente*, in «Noûs», vol. 29, n. 1, 1995, pp. 21-45; e in D. Pereboom, *Living Without Free Will*, Cambridge, Cambridge University Press, 2001.

¹⁴ Il determinismo è la tesi (metafisica) per cui tutti gli eventi del mondo sono conseguenza senza eccezioni di eventi precedenti in coerenza con le leggi fisiche immutabili, in una concatenazione che parte dal primo evento identificabile al futuro più lontano. Anche gli esseri umani sono, per definizione, soggetti al determinismo fisico.

arbitrio e la responsabilità morale sono soltanto illusioni, il concetto di imputabilità perde inevitabilmente di senso, e la sanzione tradizionale – che dovrebbe proporsi come provvedimento ragionevole e proporzionale – rischia di assumere i connotati di una “pena crudele e inusuale” contraria ai principi di giustizia e di umanità cui dovrebbe ispirarsi.

Il determinismo neuroscientifico solleva poi un ordine di questioni cosiddette interne. Per esempio, alla luce delle nuove conoscenze sullo sviluppo cerebrale, che prosegue fino a 21-22 anni e oltre nella corteccia prefrontale (un’area fortemente coinvolta nel controllo dei comportamenti impulsivi) si è infatti ridiscusso il concetto legale di “maggiore età” e la soglia della sua fissazione. Un’età convenzionale, fissata solo sulla base di osservazioni prescientifiche, non sembra corrispondere a stati di cose reali, poiché le neuroscienze ci presentano un quadro più sfumato, in cui le nozioni di normalità e patologia, minore e maggiore età (cioè una fase di maturazione del cervello sufficiente a dare il controllo degli impulsi istintivi) rappresentano gli estremi di un continuum biologico graduale. Le nuove conoscenze in questo ambito, oltre a produrre un attrito potenziale tra dati neuroscientifici e categorie giuridiche, hanno già prodotto in Olanda una riconsiderazione dei requisiti per essere ricompresi nella sfera della giustizia minorile¹⁵.

2.2. “Trattare” il crimine

L’argomentazione degli *incompatibilisti forti* insiste su un mutamento di prospettiva in merito alla giustificazione della pena: se non siamo davvero liberi, o perlomeno non lo siamo nella misura in cui ritenevamo di esserlo, allora la pena non può più essere giustificata sulla base del concetto di merito (*desert*). Dal momento che – molto semplificando – sarebbero i geni difettosi e i cervelli malfunzionanti i reali “colpevoli” della cattiva condotta e non l’individuo consapevole nella sua interezza, risulta fallace la concezione secondo cui gli esseri umani meriterebbero di essere rimproverati (o lodati) e puniti (o premiati).

Le ormai note sentenze di Trieste (2007)¹⁶ e Como (2011)¹⁷ costituiscono due esempi italiani di applicazione pratica di un simile approccio

¹⁵ Sul caso olandese, cfr. S. Schlem, *Real Neurolaw in the Netherlands: The Role of the Developing Brain in the New Adolescent Criminal Law*, in «Frontiers in Psychology», vol. 11, n. 1762, 2020, pp. 1-5; e A. Lavazza e F. Corso, *Neurolaw and the Mind-Brain Problem in Practice: The Case of Psychological Immaturity and Brain Immaturity*, in «Teoria e critica della regolazione sociale», vol. 1, n. 22, 2021, pp. 39-59.

¹⁶ Sentenza della Corte d’Assise d’Appello di Trieste n. 5/2009.

¹⁷ Sentenza del Tribunale di Como n. 536/2011.

neurogiuridico; in entrambi i casi, infatti, e nonostante la gravità dei reati (omicidio e tentato omicidio), la sanzione penale fu significativamente ridotta a seguito di un'attenta valutazione delle caratteristiche genetiche e psichiatriche degli imputati, da cui deriva – secondo i periti della difesa, le cui prospettazioni vennero accolte dai giudici – una seria compromissione della capacità di autocontrollo degli impulsi violenti.

Seppur a uno stadio ancora preliminare, queste due sentenze mostrano come una ponderata considerazione dell'approccio neuroscientifico latamente determinista imponga (sempre secondo l'interpretazione radicale) una mutazione teorica e pratica del diritto. In particolare, si fa strada un paradigma alternativo, in cui le nozioni di reato e pena assumono un significato diverso rispetto a quello tradizionale.

Da un lato la prospettiva neuroscientifica tende a considerare il delitto come la manifestazione (il sintomo) della condizione patologica dell'autore di reato, mentre quest'ultimo per il sistema penale classico rappresenta – ad alto livello di astrazione – la rottura di un equilibrio che deve essere ripristinato attraverso la sanzione. Dall'altro lato, e come conseguenza diretta della patologizzazione del comportamento deviante, la pena viene progressivamente trasformata in due sensi: sul piano individuale, non è più la sanzione per la scelta di delinquere ma deve essere l'“antidoto” alla malattia soggiacente le condotte illecite; sul piano collettivo, va intesa come quell'insieme di misure preposte alla salvaguardia dell'incolumità pubblica: rientra cioè in un approccio consequenzialista, che non valuta quanto compiuto ma si concentra sugli effetti e le circostanze.

In senso generale, si tratta di un punto di vista che mira a invertire la direzione dello sguardo della giustizia: a essere decisiva non è più l'esigenza di punire un reato commesso *nel passato*, ma la necessità di evitare che i crimini si verifichino nuovamente *in futuro*, cosicché la pena può e deve essere giustificata solamente sulla base dei risultati benefici che essa può produrre a livello sociale. Nella prospettiva del consequenzialismo neuroscientifico, il crimine dovrebbe essere “trattato” (più che punito) mediante un insieme di misure (anche preventive) atte a tutelare il reo e la collettività.

A titolo esemplificativo, è possibile suddividere queste misure in due ampie categorie: 1) quelle rientranti nel filone dei *neurointerventions*¹⁸, ossia quell'insieme di interventi biomedici (tra cui la somministrazione di farmaci serotoninergici) atti a ridurre l'aggressività, ai quali il soggetto che ha compiuto (o potrebbe compiere) atti criminali può sottoporsi in

¹⁸ Si veda, ad esempio, J. Ryberg, *Neurointerventions, Crime and Punishment. Ethical Considerations*, Oxford, Oxford University Press, 2020.

cambio di uno sconto della pena o in alternativa al carcere; 2) quelle che si concentrano sulla limitazione *preventiva* delle libertà personali del soggetto considerato *potenzialmente* violento ai fini di salvaguardare la sicurezza pubblica (e personale). In quest'ottica si inserisce, ad esempio, il modello della quarantena (anche preventiva) avanzato da Pereboom¹⁹, e ripreso successivamente da Caruso²⁰.

Non si intende in questa sede entrare ulteriormente nel merito del dibattito inerente le complesse problematiche etiche che tali proposte hanno comprensibilmente alimentato. Basti qui evidenziare che questo modello (neuro)conseguenzialista, pur avendo il merito di perseguire ideali umanitari fondamentali attraverso un metodo che sia il più possibile fondato sulle acquisizioni scientifiche, non è esente da criticità e lacune.

Nello specifico, si possono sollevare critiche di carattere generale alla posizione *radicale* vista precedentemente in merito alle premesse scientifiche e teoriche dell'incompatibilismo forte, nonché all'applicazione pratica di tali presupposti in sede penale, perplessità che spingono verso la posizione detta *moderata*.

3. La realtà giuridica e del senso comune

Nella sezione precedente, si è messo in luce come la realtà neuroscientifica risulti controintuitiva e perfino disturbante, se rapportata alla visione implicita nel senso comune e nel diritto. In primo luogo, la credenza nel libero arbitrio e nella responsabilità morale è talmente radicata nelle nostre pratiche che stenta a essere scalfita in maniera significativa dagli esiti sperimentali delle neuroscienze.

D'altronde, è l'impalcatura stessa del diritto penale a riflettere, esprimendola in codici e sanzioni, la forte convinzione (che seppure fosse illusoria viene percepita come reale) di essere gli artefici del proprio destino e, dunque, responsabili delle proprie azioni. Sul piano meramente fenomenologico, non tendiamo a percepire noi stessi come macchine pre-programmate e inconsapevoli, ma sentiamo, al contrario, di avere la padronanza di decisioni e comportamenti.

L'impostazione giuridica muove infatti dall'assunto fondamentale

¹⁹ Cfr. D. Pereboom, *Free Will, Agency and Meaning in Life*, Oxford, Oxford University Press, 2014.

²⁰ Cfr. G. D. Caruso, *Free Will Skepticism and Criminal Behavior: A Public Health-Quarantine Model*, in «Southwest Philosophy Review», vol. 32, n. 1, 2016, pp. 25-48.

della psicologia del senso comune secondo cui gli esseri umani sono (in linea di massima e salvo eccezioni) agenti autocoscienti, liberi e razionali a cui deve essere attribuita responsabilità morale e giuridica²¹. In altri termini, le azioni umane, ben lungi dall'essere la proiezione di meccanismi neuronali inconsci che sfuggono al vaglio della coscienza e della ragione, vengono considerate il frutto di una volontà libera e razionale. Secondo questa prospettiva, non solo si ha valido motivo di credere che gli agenti siano in grado di agire liberamente: ma è proprio questa stessa capacità di autodeterminazione a giustificare l'idea che il soggetto sia moralmente responsabile e imputabile.

La presenza di una *mens rea* che non sia riducibile agli stati cerebrali dell'agente costituisce difatti il caposaldo imprescindibile dell'elemento soggettivo del reato, comportando altresì una tacita ammissione di una sorta di dualismo mente-corpo: a essere colpevole è "l'io stabile e unitario" che si presenta capace di non soggiacere totalmente agli innegabili condizionamenti biologici. In assenza di una mente intesa classicamente come sede del pensiero cosciente e razionale, non solo risulta priva di legittimità qualsiasi forma di biasimo (responsabilità soggettiva), ma non è neppure possibile rinvenire alcun nesso psichico tra la condotta materiale e il suo autore (responsabilità oggettiva)²².

Il sistema penale classico sembra dunque enfatizzare presupposti filosofici e scientifici antitetici rispetto a quelli suggeriti dal neurodiritto. Abbiamo da un lato una mente che è di norma integra e solida, la cui realtà rende valido il concetto di merito e biasimo; d'altro canto, la prospettiva basata sulle neuroscienze rifiuta questa premessa su basi empiriche. L'immagine dell'essere umano che il senso comune e il diritto hanno contribuito a formulare è (o sembra essere) scientificamente falsa, cosicché l'impianto giuridico si trova a dovere fare fronte a risultanze neuroscientifiche incompatibili con il suo impianto.

In questo senso, si possono evidenziare due strategie di "resistenza". La prima, detta *compatibilista*, mira a "neutralizzare" il potenziale impatto pragmatico delle neuroscienze sul diritto penale, salvando al contempo determinismo e libero arbitrio; la seconda, che potremmo chiamare *isolazionista*, si propone invece di rifiutare lo sguardo totalmente naturalistico, propendendo per una netta separazione tra il discorso descrittivo e quello normativo.

²¹ Per psicologia del senso comune si intende un insieme sistematizzato dagli studiosi di intuizioni e credenze diffuse, che ci vengono probabilmente dall'evoluzione e che sono "cablate" nel nostro cervello.

²² Cfr. A. Lavazza e L. Sammicheli, *Il delitto del cervello*, cit., p. 40.

Ovviamente, queste strategie muovono dalla constatazione che la scienza abbia realmente portato elementi che possono aumentare lo scetticismo verso l'esistenza del libero arbitrio e, di conseguenza, compiono mosse concettuali per sterilizzare questa "minaccia" alle basi del diritto. Esiste, tuttavia, anche una robusta critica sul piano scientifico agli esperimenti "alla Libet" e alla loro interpretazione, la quale sostiene che nulla di definitivo è stato dimostrato in merito alla nostra capacità di scelta libera e che il dibattito secolare resta aperto²³.

La discussione nel campo strettamente scientifico e il rapido mutare delle convinzioni dei neuroscienziati in base ad esperimenti che si susseguono in tempi rapidi dovrebbe costituire un elemento di cautela rispetto ad estrapolazioni totali e immediate dal laboratorio alle aule di giustizia e, addirittura, ai codici. In questo senso, i tempi lunghi del diritto sono un utile correttivo a una certa fretta delle pretese avanzate dal neurodiritto. Nello stesso, è importante continuare a essere ricettivi nei confronti della scienza, per non alzare muri che danneggiano il diritto stesso e, soprattutto, coloro che vi si affidano.

3.1. *Il compatibilismo*

Il compatibilismo attuale circa il libero arbitrio si rifà a quello classico sostenuto da filosofi come Hobbes²⁴, Locke²⁵ e Hume²⁶. Ad essere centrale nel compatibilismo tradizionale è l'idea che il determinismo filosofico sia

²³ Si vedano A. Mele, *Liberi! Perché la scienza non ha confutato il libero arbitrio*, Roma, Carocci, 2015; A. Lavazza, *Il libero arbitrio tra scienza e neuroetica*, in «Notizie di Politeia», vol. XXXVII, n. 142, 2021, pp. 109-113; M. De Caro, A. Lavazza, G. Sartori (a cura di), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, Torino, Codice Edizioni, 2019 (2ª ed.).

²⁴ «Dall'uso delle parole libero arbitrio non si può inferire alcuna libertà della volontà, del desiderio o dell'inclinazione, ma [quella che si può inferire è la libertà dell'uomo, la quale consiste nel non incontrare arresti nel fare ciò che ha la volontà, il desiderio o l'inclinazione di fare], T. Hobbes, *Leviatano*, Roma; Bari, Laterza 2008, p. 176.

²⁵ «Questo, dunque, è evidente che un uomo non è libero di volere o non volere una cosa qualunque che sia in suo potere, una volta che l'abbia considerata: poiché la libertà consiste nel potere di agire o di omettere l'azione, e in questo soltanto.», J. Locke, *Saggio sull'intelligenza umana*, Roma; Bari, Laterza, 2011, p. 265.

²⁶ «Con libertà, dunque, vogliamo significare soltanto un potere di agire o di non agire secondo le determinazioni della volontà; ossia che, se preferiamo restar fermi, possiamo; se preferiamo muoverci, egualmente possiamo. Ora questa ipotetica libertà tutti ammettono che appartenga a ognuno che non sia prigioniero e incatenato.» D. Hume, *Ricerca sull'intelletto umano*, Roma; Bari, Laterza, 1996, p. 101.

non soltanto compatibile con la libertà umana, ma che ne costituisca la condizione necessaria. I compatibilisti classici ritengono infatti che l'unica libertà concepibile sia quella di agire in conformità con la propria volontà; quest'ultima, tuttavia, risulta sempre determinata da fattori al di là del controllo dell'individuo. Contrariamente a quanto predicato dai libertari, la volontà non può mai dirsi effettivamente libera per via della condizione ontologica stessa degli esseri umani che, in quanto organismi viventi, non possono che risultare determinati, innanzitutto dalla propria finitezza²⁷. Ad essere libera è l'azione, la possibilità di agire (o non agire) secondo volontà.

La tesi dei compatibilisti classici è stata in seguito ripresa e aggiornata dai sostenitori del compatibilismo contemporaneo. Caratteristica di questa prospettiva è la posizione scettica nei confronti delle presunte implicazioni antropologiche e giuridiche del riduzionismo neuroscientifico. Uno degli esempi più significativi di compatibilismo contemporaneo è quello di Morse; per quest'ultimo, il fatto che le nostre azioni discendano da una concatenazione di cause è irrilevante ai fini dell'imputabilità.²⁸

In un universo causalmente deterministico, tutti i fenomeni, comprese le azioni umane, sono completamente causati. Se la causazione fosse di per sé una condizione scusante, nessuno potrebbe mai essere responsabile di alcunché. Pertanto, la causazione non può essere una condizione scusante per la legge e la morale, che considerano alcune persone responsabili e ne scusano altre.

Con il compatibilismo contemporaneo, prende piede l'idea secondo cui la libertà umana non debba intendersi come una facoltà trascendente rispetto alla realtà materiale; al contrario, essa si inserisce pienamente nel paradigma naturalistico, costituendo un prodotto e una condizione dell'evoluzione naturale²⁹.

Questa prospettiva naturalistica del libero arbitrio (e, in generale, della mente umana) intende pertanto confutare la tesi incompatibilista fin dai suoi presupposti, evidenziando come la presunta dicotomia libertà/determinismo – con tutte le sue potenziali ripercussioni pratiche sull'imputazione della responsabilità e sulla sanzione penale – costituisca in

²⁷ «In questa prospettiva, non è tanto che siamo liberi anche se siamo determinati; siamo liberi grazie al fatto che siamo determinati», M. De Caro, *Il libero arbitrio. Una introduzione*, Roma, Laterza 2004, p. 61.

²⁸ S. J. Morse, *The Non-Problem of Free Will in Forensic Psychiatry and Psychology*, in «Behavioral Sciences & the Law», vol. 25, n. 2, pp. 203-220, pp. 216-217 (traduzione degli autori).

²⁹ Cfr. D. C. Dennett, *Freedom Evolves*, Viking, New York; trad. it. *L'evoluzione della libertà*, Raffaello Cortina, Milano, 2004.

realtà un falso problema: possiamo continuare a ritenerci responsabili delle nostre azioni anche in un universo determinista.

Ora, al di là delle critiche degli incompatibilisti nei confronti del compatibilismo, si può tuttavia individuare un terreno comune su cui poggiano le due posizioni; sia i compatibilisti sia gli incompatibilisti, pur giungendo a conclusioni opposte, maturano le proprie argomentazioni rimanendo ancorati alla prospettiva naturalistica. Per i primi, il modo in cui l'evoluzione modella le inclinazioni, le emozioni e le facoltà morali rende queste caratteristiche umane ancora più reali e concrete; per gli altri, il fatto che "il mondo sia governato da una concatenazione immutabile di eventi"³⁰ implica ammettere l'illusorietà del libero arbitrio.

Alcuni studiosi hanno tuttavia provato a ridimensionare il potenziale rivoluzionario sul diritto delle neuroscienze avanzando argomentazioni di chiaro stampo anti-naturalistico. Su questo versante si collocano coloro che ritengono sia necessario recuperare la distinzione tracciata da Hume tra piano descrittivo (che pertiene alla scienza) e piano normativo (che pertiene al diritto) evitando così la confusione tra i due ambiti di fatto incommensurabili.

3.2. *La tesi antinaturalistica*

Finora si è potuto osservare come gran parte del dibattito neurogiuridico si inserisca nel quadro della "naturalizzazione del pensiero come tesi metafisica"³¹. L'assunto su cui si basa tale impostazione è che tutti i fenomeni – inclusi quelli specificatamente umani come le emozioni morali e il possesso dei concetti di giusto e sbagliato – siano in linea di massima riconducibili alle leggi del mondo naturale.

Così, ad esempio, scrive Nunziante:

C'è *prima* l'ontologia fisicalista, c'è prima il modello epistemico delle scienze naturali, c'è *prima* il dato osservativo rispetto all'enunciato teorico, c'è *prima* l'epistemologia rispetto alla filosofia, ci sono *prima* le neuroscienze e così via. Ma in tutto questo è come se si producesse una sorta di gigantesca "metafisica" di ritorno.³²

³⁰ A. Lavazza e L. Sammiceli, *Il delitto del cervello*, cit., p. 80.

³¹ A. M. Nunziante, *La naturalizzazione del pensiero come tesi metafisica*, in S. Fuselli (a cura di), *Neurodiritto. Prospettive epistemologiche, antropologiche e biogiuridiche*, Milano; Udine, Mimesis edizioni 2016, cit., p. 17.

³² *Ivi*, cit., p. 51.

Si possono tuttavia rilevare alcuni approcci caratterizzati da un maggiore scetticismo rispetto alla proposta di naturalizzare la morale e il diritto. A questo proposito, e in via preliminare, vanno rimarcate due cautele relative ai dati neuroscientifici in sé; in primo luogo, occorre riconoscere che le neuroscienze costituiscono un campo di studi ancora relativamente acerbo; in secondo luogo, i dati delle neuroscienze sembrano indicare come il rapporto cervello-comportamento debba essere considerato in termini più propriamente probabilistici, anziché deterministici.

Esiste poi una corrente di pensiero *isolazionista* che, al di là dei giudizi sull'affidabilità delle risultanze neuroscientifiche, sostiene l'importanza di mantenere la distinzione tra piano normativo e piano descrittivo. Secondo questo approccio, i concetti della filosofia e del diritto possiedono uno statuto epistemologico indipendente da quello delle neuroscienze, per cui risulterebbe logicamente impossibile verificare un concetto normativo con gli strumenti della scienza.

In *Philosophical Foundations of Neuroscience*³³, Bennett e Hacker riprendono efficacemente questo dibattito, distinguendo funzioni e campi di pertinenza degli studi empirici e dell'argomentazione concettuale. Così, secondo gli autori, le neuroscienze rispondono a quesiti di natura meramente sperimentale circa il sistema nervoso e il modo in cui il suo funzionamento rende possibile la consapevolezza, la volizione e le emozioni morali. Su un altro versante, invece, non sono indagabili scientificamente le questioni concettuali, poiché queste non sarebbero né vere né false:

Le questioni concettuali precedono le questioni di verità e falsità. Sono questioni che riguardano le nostre *forme di rappresentazione*, non questioni che riguardano la verità o la falsità delle affermazioni empiriche. Queste forme sono presupposte da affermazioni scientifiche vere (*e* false) e da teorie scientifiche corrette (*e* sbagliate). Esse non determinano ciò che è empiricamente vero o falso, ma piuttosto ciò che ha o non ha senso. Pertanto, le questioni concettuali non sono suscettibili di indagine e sperimentazione o di teorizzazione scientifica. Infatti, i concetti e le relazioni concettuali in questione sono *presupposti* da qualsiasi indagine e teorizzazione.³⁴

Bennett e Hacker definiscono “fallacia mereologica” il tentativo dei neuroscienziati di attribuire al cervello caratteristiche psicologiche; in realtà, la questione che si pone in merito al rapporto mente-cervello

³³ M. R. Bennett e P. M. S. Hacker, *Philosophical Foundations of Neuroscience*, John Wiley & Sons, (2^a edition) 2022.

³⁴ *Ivi*, cit., pp. 1-2 (traduzione degli autori).

(ovvero se le caratteristiche psicologiche possano o meno essere attribuibili al funzionamento cerebrale) altro non sarebbe che una questione filosofica o concettuale, non scientifica.³⁵

È bene precisare che queste considerazioni non si pongono l'obiettivo di riproporre una sorta di dualismo cartesiano rivisitato in chiave contemporanea; non si tratta di ripristinare una separazione *ontologica* tra facoltà mentali e processi cerebrali, quanto di rilevare la distinzione epistemologica tra verità di ragione e verità di fatto, tra realtà concettuale e realtà empirica.

3.3. *Il ruolo delle emozioni*

Un diverso approccio critico nei confronti dei sostenitori radicali della rivoluzione neurogiuridica mira a recuperare l'importanza della dimensione emotiva nei processi decisionali. Si tratta, più che altro, di una critica diretta ai risvolti pragmatici (più che teorici) dell'*hard incompatibilism* sul diritto, ossia quelli relativi, nello specifico, alla sanzione penale.

Come è già stato osservato, le istanze del determinismo neuroscientifico propendono per uno slittamento del significato e della funzione della pena in virtù dell'impossibilità di considerarci effettivamente liberi e responsabili. Da qui, si fa strada la pretesa dell'abbandono definitivo del retributivismo, che proprio sui concetti di libero arbitrio e responsabilità si fonda.

Occorre tuttavia osservare che questa prospettiva radicale non sembra rendere sufficientemente conto delle nostre naturali inclinazioni punitive nei confronti di chi reca danno al prossimo. La dimensione emotiva degli esseri umani, ampiamente trascurata nel diritto, viene ad esempio riabilitata da Strawson. La tesi centrale del suo celebre articolo *Freedom and Resentment*³⁶ è che i nostri "atteggiamenti reattivi" (*reactive attitudes*) – tra cui l'indignazione, il risentimento e il perdono – anziché essere un elemento d'intralcio quando si tratta di attribuire o meno un certo grado di responsabilità agli individui, costituiscono in realtà il perno attorno a cui ruota il concetto stesso di responsabilità.

In questo modo, la nozione di responsabilità avanzata da Strawson si proporrebbe come "terza via"³⁷ rispetto a quella implicita nel modello neurogiuridico radicale – in quanto essa sarebbe innanzitutto connotata

³⁵ Cfr. *ivi*, pp. 82-84.

³⁶ P. F. Strawson, *Freedom and Resentment*, in «Proceedings of the British Academy», vol. 48, 1962, pp. 1-25.

³⁷ M. De Caro, *Il libero arbitrio*, cit., p. 108.

socialmente, e non andrebbe dunque interpretata alla stregua di un qualcosa di oggettivo e impersonale – nonché rispetto alla concezione tradizionale della responsabilità come imprescindibile dal concetto di libero arbitrio, dal momento che:

Secondo Strawson, questa concezione ricerca un'impossibile giustificazione *esterna* ai sentimenti morali reattivi che sottostanno alle attribuzioni di responsabilità, laddove sono proprio tali sentimenti che costituiscono o meglio *esprimono* la responsabilità, al punto che – al contrario di ciò che pensano i libertari – noi non potremmo smettere di considerare i nostri simili responsabili per le loro azioni nemmeno nel caso in cui disponessimo di valide ragioni per credere nella verità del determinismo causale.³⁸

I nostri atteggiamenti reattivi si sarebbero dunque fissati e calibrati evolutivamente in maniera tale da attivarsi in risposta a condotte che si reputano intenzionali e a disattivarsi in risposta a condotte che si reputano accidentali, cosicché ad assumere un ruolo centrale sarebbe la percezione diretta di “un comportamento intenzionale come *in sé rinviante a un significato*”³⁹.

In questo senso, le emozioni non sono più *solamente* meri stati affettivi prodotti da cambiamenti neurofisiologici che si verificano a seguito della percezione di uno stimolo⁴⁰, ma sono *anche* stati cognitivi in cui è implicito un giudizio come contenuto semantico. John Deigh definisce “neo-jamesiani”⁴¹ coloro che recuperano la prospettiva di James, integrandola con la teoria cognitivista delle emozioni:

I neo-jamesiani, come James, ritengono che i sentimenti in cui consistono le emozioni siano i sentimenti dei cambiamenti corporei che si verificano durante un episodio di emozione. A differenza di James, però, i neo-jamesiani ritengono che questi sentimenti non si limitino a registrare i cambiamenti corporei. In realtà, essi sono

³⁸ *Ivi*, cit., p. 109.

³⁹ A. Lavazza e L. Sammiceli, *Se non siamo liberi, possiamo essere puniti?*, in M. De Caro, A. Lavazza e G. Sartori (a cura di), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, cit., p. 163.

⁴⁰ È questa la concezione jamesiana delle emozioni, ritenute semplici epifenomeni di cambiamenti corporei e prive di forza motivazionale. Contrariamente alla psicologia del senso comune, James ritiene infatti che le emozioni non causino il cominciamento dell'azione. Cfr. W. James, *The Principles of Psychology*, vol. 1 e 2, New York, Dover, 1950.

⁴¹ Tra i “neo-jamesiani”, uno degli autori più rappresentativi è senz'altro Antonio Damasio. Cfr. A. Damasio, *Descartes' Error*, New York, Avon Books, 1995.

anche portatori di un significato.⁴²

Seguendo questa prospettiva teorica, si può dunque sostenere che le nostre emozioni retributive, lungi dal contrapporsi all'istanza neuroscientifica neutra, sono di fatto intrise di un valore espressivo e semantico che non sembra possa essere ragionevolmente ignorato. Inoltre, le emozioni sembrano influenzare il comportamento anche quando non sono percepite a livello consapevole, come sta mostrando un filone di ricerca che indaga i fenomeni legati all'inconscio cognitivo, dove processi motivazionali sono all'opera anche se non ce ne rendiamo conto.

Questo apre alla considerazione di un settore specifico d'indagine sui moventi del comportamento che potrebbe avere importanti ricadute sul diritto, dato che non è basato su controversi esperimenti recenti delle neuroscienze ma su apparentemente più semplici studi di psicologia empirica. Tali studi forniscono uno spaccato importante sulla fragilità delle ragioni del nostro agire e sullo scarso controllo che abbiamo su di esse a causa di influenze inconsapevoli provenienti dall'ambiente. Tutto ciò sembra indicare che molto più spesso di quanto pensiamo siamo vittime di "manipolazioni" di cui non ci rendiamo conto.

4. *Il situazionismo*

Il comportamento umano pare influenzato sia da fattori interni sia da fattori esterni. Le neuroscienze cognitive tendono a focalizzarsi sui fattori interni; il cosiddetto situazionismo, invece, si concentra sulle modalità attraverso cui stimoli ambientali di vario tipo possono orientare inconsciamente l'individuo, indirizzandone le scelte e i comportamenti.

Secondo i situazionisti⁴³, le nostre decisioni sono in larga parte influenzate da processi automatici e inconsci che vengono innescati da circostanze ambientali (secondo l'effetto *priming* per cui stimoli di cui non siamo consapevoli mettono in azione processi cognitivi e motivazionali inconsci). Le azioni umane non sarebbero dunque il frutto della volontà

⁴² J. Deigh, *Concepts of Emotions in Modern Philosophy and Psychology*, in P. Goldie (ed.), *The Oxford Handbook of Philosophy of Emotion*, Oxford, Oxford University Press 2010, p. 32 (la traduzione è degli autori).

⁴³ Cfr. J. M. Doris, *Lack of Character: Personality and Moral Behavior*, Cambridge, Cambridge University Press, 2002; e K. A. Appiah, *Experiments in Ethics*, Cambridge, MA, Harvard University Press, 2008.

consapevole degli individui sulla base di ragioni, bensì il risultato di conseguenze automatiche di fattori esterni capaci di condizionarci⁴⁴.

In questa prospettiva, più che i tratti della personalità, risulterebbero essenziali gli “input situazionali”, gli stimoli esterni che orientano l’azione senza la nostra partecipazione attiva (quasi fossimo sotto ipnosi). Attraverso lo studio approfondito delle situazioni e dei relativi “inneschi”, sarebbe quindi possibile sia manipolare sia prevedere il comportamento degli individui.

Il situazionismo si avvale di una mole notevole di dati sperimentali. Ad esempio, in un esperimento condotto da John Bargh e collaboratori nel 1996⁴⁵, e divenuto noto per il suo essere piuttosto sorprendente, si era dimostrato come gli stimoli esterni avessero un effetto sul ritmo di movimento dei soggetti. Nello studio, a due gruppi di studenti volontari americani venne chiesto di comporre alcune frasi usando una serie di parole messe a loro disposizione, ma solo ad un gruppo vennero affidate parole che rievocavano lo “stereotipo” della vecchiaia (come “rughe”, “saggio”, “conservatore”, “Florida”). Al termine del compito, i gruppi dovevano recarsi in un’altra stanza. Il risultato fu che il gruppo che era stato esposto alle parole relative alla vecchiaia camminava molto più lentamente rispetto al gruppo di controllo, a dimostrazione, si presumeva, del condizionamento implicito esercitato dal compito.

Altri esperimenti hanno avuto l’obiettivo di studiare fenomeni sociali come il conformismo e l’obbedienza all’autorità. Ben noti, in tal senso, sono i risultati degli esperimenti di Asch (1951)⁴⁶ e Milgram (1961)⁴⁷. Il primo si proponeva di analizzare l’“effetto maggioranza” sul giudizio individuale, e rivelò come la maggior parte dei soggetti mostrasse una tendenza conformistica quando esposta alla pressione del gruppo di riferimento. Nell’esperimento di Milgram, invece, si intendeva studiare in che modo gli individui rispondessero agli ordini delle autorità, anche quando questi compiti implicassero azioni immorali e crudeli (come somministrare scariche elettriche a un altro essere umano: finte ma credute reali dai partecipanti). La maggior parte dei soggetti, pur riluttante, non esitò di fronte ai comandi e somministrò le scosse ai compagni che

⁴⁴ Cfr. A. Lavazza, *Why Cognitive Sciences Do Not Prove That Free Will Is an Epiphenomenon*, in «Frontiers in Psychology», vol. 10, n. 326, 2019, pp. 1-11.

⁴⁵ J. A. Bargh, M. Chen e L. Burrows, *Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action*, in «Journal of Personality and Social Psychology», vol. 71, n. 2, 1996, pp. 230-244.

⁴⁶ S. E. Asch, *Effects of Group Pressure upon the Modification and Distortion of Judgments*, in R. W. Allen, L. W. Porter e H. L. Angle (ed.), *Organizational Influence Processes*, Routledge, 2016.

⁴⁷ S. Milgram, *Obedience to Authority*, New York, Harper and Row, 1969.

fingevano di contorcersi dal dolore.

In linea generale, questi esperimenti sembrano confermare l'idea secondo cui gli individui sono costantemente immersi in un flusso ininterrotto di stimoli (situazionali, ambientali, sociali) che inducono una risposta comportamentale inconsapevole. Pertanto, sia il determinismo neuroscientifico radicale che il situazionismo mettono in luce il ruolo fondamentale della dimensione non consapevole nel processo decisionale. Un elemento che riduce fortemente la responsabilità individuale rispetto alle proprie scelte.

Pur non negando l'importanza dei fattori esterni, le neuroscienze fanno principalmente perno sulle caratteristiche interne dell'individuo (geni, chimica cerebrale); con il situazionismo, invece, la prospettiva viene in qualche modo rovesciata, cosicché a risultare determinanti per scoprire i meccanismi alla base delle scelte umane sarebbero soprattutto i fattori esterni. Il situazionismo, in netto contrasto con la psicologia del senso comune, svela uno scenario che pare smentire l'idea secondo cui gli esseri umani sono agenti razionali guidati nel comportamento dai propri valori e dalle proprie preferenze in modo costante e coerente.

Va tuttavia sottolineato che l'esistenza di meccanismi automatici inconsci di risposta a fattori ambientali e situazionali, di per sé, non comporta necessariamente l'irrazionalità di tali moventi. Si può per esempio parlare di "scorciatoie cognitive" che si attivano in determinate circostanze, bypassando l'agentività cosciente, che richiede sforzi e tempi di risposta più lunghi. Così, ad esempio, Vargas:

Quello che chiamiamo "istinto" è (almeno a volte) il modo in cui Madre Natura codifica una sorta di razionalità limitata nei nostri meccanismi di base. Lo consideriamo un semplice istinto solo quando scopriamo i limiti di questa razionalità o quando essa entra in conflitto con i nostri obiettivi di ordine superiore.⁴⁸

Quello che però risulta più rilevante è la cosiddetta crisi di replicazione, che ha colpito soprattutto la psicologia empirica e numerosi esperimenti su cui si basa il situazionismo. La crisi della replicazione (detta anche della replicabilità e della riproducibilità) nasce dalla scoperta che gli esiti di molti studi scientifici sono risultati difficili o impossibili da riprodurre. Poiché la riproducibilità dei dati empirici è una parte essenziale del

⁴⁸ M. Vargas, *Situationism and Moral Responsibility: Free Will in Fragments*, in A. Clark, J. Kiverstein, T. Vierkant (ed.), *Decomposing the Will*, New York, Oxford University Press, 2013 (la traduzione è degli autori).

metodo scientifico, tali fallimenti minano la credibilità delle teorie che si basano su di essi e mettono potenzialmente in discussione parti sostanziali della conoscenza scientifica. La crisi della riproducibilità ha riguardato principalmente la psicologia e la medicina, dove sono stati compiuti sforzi considerevoli per indagare nuovamente gli esiti di noti esperimenti, per determinarne l'affidabilità e, quando questa è stata messa in discussione, le ragioni del fallimento⁴⁹.

La psicologia, in particolare la psicologia sociale, e gli studi sull'effetto priming sono emersi tra i più difficili da replicare⁵⁰. Le ragioni sono complesse e oggi ampiamente indagate, riconducibili alle modalità statistiche utilizzate e agli intervalli di confidenza utilizzati. Inoltre, il tipo stesso degli esperimenti sembra indurre una sovrastima degli effetti. Altre concause, come la ricerca di nuovi risultati da parte sia delle istituzioni di ricerca sia delle riviste scientifiche, unite alla pressione a produrre dati e pubblicare in giornali ad alto *impact factor*, contribuiscono a fare sì che gli studiosi testino ipotesi di forte richiamo e non impieghino tempo e risorse a ripercorrere sentieri già battuti.

Non si tratta quindi, se non in una infinitesima parte, di deliberate manipolazioni, ma di distorsioni insite nei metodi stessi utilizzati. Lo stesso studio sugli studenti che si comportano "da anziani" si è dimostrato impossibile da replicare con gli stessi effetti⁵¹.

È altamente indicativo che il premio Nobel per l'Economia Daniel Kahneman abbia "ritrattato" il capitolo IV, intitolato "Il meccanismo associativo", del suo famosissimo libro *Pensieri lenti e veloci*⁵², perché a una riconsiderazione basata proprio sulla crisi di replicabilità, tutti gli esperimenti che cita hanno un R-Index troppo basso. L'R-Index è una misura quantitativa dell'integrità della ricerca che può essere utilizzata per valutare la replicabilità statistica di un insieme di studi. Se si va sotto il 50%, i risultati possono essere del tutto casuali.

La lezione che si può trarre è che se persino un gigante della disciplina ha preso per buoni i dati empirici su cui si è costruita l'ipotesi situazionista, pur molto recente, ciò è accaduto perché essa sembrava scientificamente

⁴⁹ M. Baker, *1,500 Scientists Lift the Lid on Reproducibility*, in «Nature», vol. 533, n. 7604, 2016, pp. 452-454.

⁵⁰ Open Science Collaboration, *Estimating the Reproducibility of Psychological Science*, in «Science», vol. 349, n. 6251, 2015.

⁵¹ R. F. Baumeister, D. M. Tice, B. J. Bushman, *A Review of Multi-site Replication Projects in Social Psychology: Is It Viable to Sustain Any Confidence in Social Psychology's Knowledge Base?*, in «Perspectives on Psychological Science», in corso di pubblicazione.

⁵² D. Kahneman, *Pensieri lenti e veloci*, Milano, Mondadori, 2012.

fondata. E scientificamente fondato significa convalidato da ricercatori autorevoli con affiliazioni a importanti università che pubblicano lavori su riviste di eccellente reputazione tramite la revisione anonima di colleghi di pari livello.

Tutto questo, quindi, basta o non basta per costruire una solida teoria scientifica? Non si può rispondere se non si introduce la variabile “tempo”. Il metodo scientifico e la sua codificazione anche pragmatica (università, riviste, ecc.) funziona ottimamente, posto che si dia un periodo sufficiente perché l'intera comunità dei ricercatori metta in atto tutte le procedure e i controlli. Non bisogna infatti dimenticare che la crisi di replicabilità è emersa nello stesso circuito attraverso il quale erano passati gli studi poi contestati.

Oggi è stata introdotta la pratica dei rapporti preregistrati. Si chiede agli autori di presentare prima della raccolta dei dati una descrizione dei metodi che verranno impiegati. Una volta che il disegno sperimentale è stato sottoposto alla revisione anonima dei pari, la pubblicazione dei risultati viene provvisoriamente garantita, in base al rispetto del protocollo preventivamente proposto. Uno degli obiettivi è proprio quello di incoraggiare la pubblicazione di studi con metodi rigorosi.

5. *Conclusione*

Sulla base di quanto detto finora, si può affermare che il situazionismo in quanto tale non è stato del tutto screditato, perché numerosi studi hanno retto alla replicazione e altri saranno condotti in futuro. Il suo status è però meno forte e, pertanto, anche la sua applicazione in ambito giudiziario sarebbe più problematica. Va ricordato che il situazionismo aveva creato un rilevante problema per la teoria metaetica delle virtù, in quanto la dimostrazione che il carattere degli individui è molto meno solido e costante di quanto si creda mette a rischio il concetto di virtù morale come tratto acquisito e stabile della persona⁵³.

Di fronte al progresso della scienza, il diritto ha dunque da considerare con attenzione l'emergere di una nuova immagine scientifica dell'essere umano, in particolare per quanto riguarda i suoi moventi e come il cervello sia la sede più rilevante per valutarli appieno. Questo però non significa

⁵³ C. L. Upton, *Virtue Ethics and Moral Psychology: The Situationism Debate*, in «The Journal of Ethics», vol. 13, n. 2, 2009, pp. 103-115.

che si debba assumere tutte le acquisizioni più recenti senza un attento vaglio critico. E, soprattutto, come il caso del situazionismo ha mostrato, il fattore tempo va considerato con particolare attenzione. D'altra parte, è quello che stabiliscono anche le linee guida per l'assunzione delle prove scientifiche nel processo: i dati e le teorie devono essere ben consolidati, proprio perché la scienza è in continuo progresso e raffinamento⁵⁴. Non si può aspettare troppo, tuttavia. Come diceva Benedetto Croce in altro contesto, se si persegue l'*infinitum perfectionis*, in una vita intera non si pubblicherà mai nulla.

Il diritto deve dunque andare alla ricerca di un fecondo dialogo interdisciplinare per aggiornarsi e, se del caso trasformarsi, alla luce di una migliore comprensione del funzionamento dell'essere umano, che non può che rimanere la sua stella polare.

⁵⁴ Per esempio, il cosiddetto *Daubert Standard* negli Stati Uniti: cfr. https://en.wikipedia.org/wiki/Daubert_standard.

Retributivismo, utilitarismo e libero arbitrio

MARIO DE CARO

SOMMARIO: 1. Utilitarismo radicale in teoria della pena – 2. La sfida deterministica – 3. La sfida dell'epifenomenismo – 4. La teoria della pena tra utilitarismo e retributivismo.

In questo saggio mi concentrerò sulla sfida lanciata oggi, su base scientifica, contro l'idea tradizionale di responsabilità. Tale sfida viene oggi declinata in due diverse maniere: secondo una versione che riprende posizioni già sviluppate in passato¹, aggiornandole alla luce dei nuovi dati, si sostiene che il quadro deterministico dei meccanismi neurofisiologici della decisione umana falsifica completamente la tradizionale concezione di *mens rea*; in una versione più radicale, si sostiene invece che la mente umana non è mai in controllo cosciente e razionale delle proprie azioni. In entrambi i casi, viene perorato l'abbandono di ogni elemento retributivistico insito nelle concezioni della pena e l'approdo a sistemi puramente utilitaristici.

1. Utilitarismo radicale in teoria della pena

Già al livello intuitivo, l'istituto della pena ha la funzione di riabilitare il condannato, di scoraggiare altri potenziali criminali e di difendere la società dagli individui pericolosi: queste sono giustificazioni di tipo *utilitaristico*, perché guardano all'utilità della pena rispetto alla società nel suo complesso. Dall'altra parte, però, abbiamo anche una nitida intuizione di diverso genere: ossia che la pena serve a ristabilire l'equilibrio della giustizia, infranto da chi si è reso responsabile di un crimine e che, dunque, costui meriti di essere punito, quali che siano le conseguenze della pena. Questa è la concezione *retributivistica* della pena, secondo cui la pena si giustifica quando un condannato la merita: e in questa prospettiva gli effetti sociali che tale inflizione potrebbe avere sono del tutto irrilevanti.

¹ C. Darrow, *Crime: Its Cause and Treatment*, Thomas Y. Crowell Company, New York, 1922.

Le concezioni utilitaristiche guardano al futuro (ossia alle conseguenze della pena), le concezioni retributivistiche guardano al passato (ossia alla colpa che il reo deve espiare). L'utilitarismo è dunque una forma di *consequenzialismo*, perché assume che per valutare la moralità delle pene si debba guardare soltanto alle loro conseguenze. Il retributivismo, invece, è una forma di *deontologismo*, in quanto ritiene che sia doveroso infliggere una pena che ristabilisca l'equilibrio della giustizia infranto dal reo.

In proposito è importante riprendere una distinzione proposta da John Mackie², secondo la quale l'ideale retributivistico può essere scomposto in due componenti, molto diverse tra loro: una *positiva* ("tutti i colpevoli meritano di essere puniti, con la necessaria severità") e una *negativa* ("solo i colpevoli meritano di essere puniti, con severità non eccessiva"). Entrambe queste componenti si incentrano sulla nozione di merito, che a sua volta presuppone quelle di responsabilità morale e, di conseguenza, quella di azione libera³: chi ha compiuto liberamente un'azione illecita, e soltanto lui, è responsabile moralmente per averla compiuta e pertanto merita di essere punito. È importante notare che, laddove la componente positiva motiva l'aspetto di severità del retributivismo (secondo cui la giustizia richiede la severa punizione di tutti i colpevoli), la componente negativa ha un connotato garantista, e ciò per due ragioni. In primo luogo, perché intima di non punire chi non lo merita – e ciò anche se tale punizione fosse potenzialmente in grado di incrementare l'utilità collettiva, come accade per esempio con le punizioni dei capri espiatori, che soddisfano la sete di vendetta della comunità. In secondo luogo, la retribuzione negativa svolge un ruolo garantista perché nega legittimità alle pene spropositate e inumane (come, ad esempio, la tortura), anche quando queste potrebbero portare ovvi benefici sociali (cosa che potrebbe accadere, per esempio, se si torturasse un terrorista in grado di confessare i futuri piani di attacco del suo gruppo).

Oggi, tuttavia, un numero crescente di scienziati, giuristi e filosofi sostiene che le neuroscienze cognitive mostrano – o, almeno, suggeriscono con forza – l'illusorietà delle idee di libero arbitrio, di responsabilità morale e di merito⁴. Se questi autori avessero ragione, le concezioni

² J. Mackie, *Persons and Values*, Clarendon Press, Oxford, 1985.

³ Sul nesso concettuale della responsabilità morale con la nozione di libero arbitrio, cfr. M. De Caro, *Il libero arbitrio. Una introduzione*, Laterza, Roma-Bari, 2021 (9ª ed.), cap. 4.

⁴ M. Gazzaniga, *The Law and Neuroscience*, in «Neuron», 40, 2008, pp. 412-415; F. Caruana, *Due problemi sull'utilizzo delle neuroscienze in giurisprudenza*, in «Sistemi intelligenti», 22, 2010, pp. 337-346; S. Harris, *Free Will*, Oxford University Press, Oxford, 2012; D. Pereboom, *Will, Agency, and Meaning in Life*, Oxford University Press, Oxford, 2014;

retributivistiche della pena (tanto nella loro componente positiva quanto in quella negativa) dovrebbero essere abbandonate e sostituite da una pura forma di utilitarismo. Secondo molti autori – come quelli appena citati – questo approdo non solo è inevitabile, ma anche auspicabile. Greene e Cohen⁵, per esempio, hanno sostenuto che, se ci convincessimo che libero arbitrio e responsabilità morale sono mere illusioni, non avrebbe alcun senso continuare a punire individui che, essendo determinati al livello genetico e neurofisiologico, non avrebbero potuto agire diversamente da come di fatto hanno agito. Ciò, tuttavia, a loro giudizio sarebbe un bene sia per gli individui puniti (almeno in cui la pena permetta loro di essere riabilitati) sia per la società nel suo complesso:

Oggi la legge tratta con fermezza, ma anche con umanità, gli individui il cui comportamento è ovviamente dovuto a forze da loro incontrollabili. Un giorno la legge potrebbe trattare in questo modo tutti i criminali condannati: ovvero li potrebbe trattare con umanità⁶.

La tesi di Greene e Cohen, e degli altri autori di questa corrente, insomma, è duplice. Primo, la scienza avrebbe ormai definito con chiarezza un quadro in cui non c'è più spazio per i fenomeni normativi e intenzionali (libero arbitrio, intenzionalità, scelta razionale ecc.), che sono indispensabili per le attribuzioni di responsabilità in senso proprio; in questa prospettiva, dunque, l'unico modo di concepire il diritto sarebbe dunque in prospettiva radicalmente utilitaristica. Secondo, questi autori sostengono che un tale mutamento, oltre ad essere giustificato, sarebbe anche profondamente auspicabile, perché renderebbe l'istituto della pena più umano. In realtà, tuttavia, disponiamo di ottime ragioni per pensare che (almeno all'attuale stato delle conoscenze) entrambe queste tesi siano erronee.

Per comprendere meglio la questione, occorre innanzi tutto considerare il tipo di sostegno empirico a cui questi autori si richiamano⁷.

G. Caruso, *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice*, Cambridge University Press, Cambridge-New York, 2021.

⁵ J. Greene, J. Cohen, *For the law, neuroscience changes nothing and everything*, in «Philosophical Transactions of the Royal Society of London B», 359, 2004, pp. 1775-85.

⁶ J. Greene, J. Cohen, *For the law*, cit. p. 1784.

⁷ Per una presentazione più articolata di questi temi, cfr. De M. Caro, A. Lavazza e G. Sartori (a cura di), *Quanto siamo responsabili?*, Codice, Torino, 2013.

2. *La sfida deterministica*

Molte delle argomentazioni contro libero arbitrio e responsabilità morale fanno riferimento alle acquisizioni della scienza, che sembrerebbero mostrare che ogni nostro comportamento è interamente determinato da fattori che sono al di là del nostro controllo⁸. In passato, a sostegno di questa tesi si faceva riferimento alla fisica (che fosse la meccanica newtoniana o la teoria della relatività) oppure a scienze sociali come la sociologia e l'antropologia (nelle loro versioni deterministiche, secondo le quali noi siamo interamente determinati dal contesto sociale e/o da quello culturale). Oggi però coloro che negano l'esistenza di libero arbitrio e responsabilità morale tendono piuttosto a richiamarsi alla gran massa di dati che giungono dalle neuroscienze, dalla genetica e dalle scienze cognitive.

Il più classico esempio in questo senso è offerto da una celebre serie di esperimenti condotti in California, a partire dagli anni Settanta, dal neurofisiologo Benjamin Libet. Il più famoso tra gli esperimenti di Libet è quello in cui si chiedeva ai soggetti sperimentali di rilassarsi e di attendere un po' prima di flettere la falange del dito indice, facendo però attenzione al momento esatto in cui veniva presa la decisione di compiere quell'azione (ossia quando avvertivano l'impulso a piegare il dito). Nel frattempo gli sperimentatori misuravano l'attività cerebrale dei soggetti tramite elettroencefalogramma. Ebbene, l'esperimento sembrava mostrare che il soggetto diveniva cosciente della propria decisione di piegare il dito un terzo di secondo *dopo* che erano cominciati i processi neurofisiologici, ovviamente inconsci (in particolare, l'attivazione del cosiddetto *potenziale di prontezza*), che causavano la flessione della falange. Molti autori hanno interpretato l'esperimento di Libet, e altri analoghi, come prova dell'illusorietà del libero arbitrio, ovvero come negazione dell'idea che ci sono casi in cui gli esseri umani possono consapevolmente determinare le azioni che compiono⁹.

⁸ Contro la nozione di responsabilità esistono anche argomentazioni a priori: cfr. M. De Caro, *Il libero arbitrio*, cit. cap. 3.

⁹ Libet medesimo, invece, offrì una differente spiegazione dei suoi esperimenti. A suo giudizio, essi potevano essere richiamati per giustificare almeno un residuo dell'idea tradizionale di libero arbitrio: ovvero la capacità, da parte della coscienza, di "porre il veto" al compimento di azioni che i processi inconsci altrimenti causerebbero. Un'obiezione sorge però naturale contro questa interpretazione (che infatti non ha molti proponenti): una volta assunto che le nostre volizioni positive abbiano determinanti neurali inconsci, perché dovremmo ritenere che le decisioni consapevoli di porre il veto a determinate

Gli esperimenti di Libet hanno inaugurato tutto un nuovo filone di ricerca teso ad indagare i processi decisionali, che ha giocato una parte di rilievo nell'ambito delle controversie su libero arbitrio e responsabilità. Il più interessante di questi studi è stato realizzato da Chun Siong Soon *et. al.*¹⁰, che hanno utilizzato una sofisticata tecnica statistica (la *pattern recognition*) per indagare l'attività cerebrale associata con una scelta apparentemente libera operata dai soggetti sperimentali. All'inizio di quell'articolo, gli autori menzionano la cosiddetta tradizione "illusionistica", secondo la quale «l'esperienza soggettiva della libertà non è nulla più di un'illusione e [...] le nostre azioni sono iniziate da processi mentali inconsci molto prima che noi diveniamo consapevoli della nostra intenzione di agire»¹¹. L'*abstract* dell'articolo mostra le ambizioni riduzionistiche che animano questo progetto:

C'è stata una lunga controversia riguardo alla possibilità che le decisioni soggettivamente "libere" siano determinate da attività cerebrali che le precedono. Noi abbiamo verificato che il risultato di una decisione può essere codificato nell'attività cerebrale della corteccia prefrontale e parietale sino a 10 sec. prima che esso raggiunga la consapevolezza. Questo ritardo presumibilmente riflette l'operazione di una rete di aree di controllo di alto livello che cominciano a preparare una decisione imminente ben prima che essa raggiunga la consapevolezza¹².

La tesi degli autori di questo articolo è che la nostra decisione conscia e solo *apparentemente* libera di compiere un'azione (si noti il termine "libere" posto tra virgolette *nell'abstract*) sia preceduta di ben 7-10 secondi dagli eventi cerebrali che determinano causalmente quell'azione (con un intervallo temporale, dunque, di più di trenta volte maggiore di quello ipotizzato da Libet). A sostegno di questa tesi, Soon *et al.* riportano i risultati di un esperimento da loro condotto, che richiede ai soggetti sperimentali di rilassarsi mentre tengono il dito indice e quello medio poggiati su due

azioni non debbano a loro volta avere determinanti neurali inconsci? Detto diversamente: che cosa mai potrebbe differenziare tanto sostanzialmente un'azione negativa da una volizione positiva, al punto che la prima si rivelerebbe l'ultimo bastione della libertà umana, mentre la seconda sarebbe interamente inconscia e dunque, secondo Libet stesso, irrimediabilmente non libera?

¹⁰ C. S. Soon-M. Brass, H. J. Heinze, J. D. Haynes, *Unconscious determinants of free decisions in the human brain*, in «Nature Neuroscience», 11, 2008, pp. 543-545.

¹¹ C. S. Soon *et al.*, *Unconscious determinants*, *op. cit.* p. 543.

¹² C. S. Soon *et al.*, *Unconscious determinants*, *idem*.

pulsanti e fissano il centro di uno schermo sul quale scorrono una serie di lettere dell'alfabeto. Viene poi richiesto loro di scegliere liberamente – quando avveriranno un impulso in tal senso – se premere il pulsante destro o quello sinistro, verificando quale lettera apparirà in quel momento sullo schermo; nel frattempo gli sperimentatori, per mezzo di una risonanza magnetica funzionale (fMRI), misurano la risposta emodinamica del cervello di questi soggetti. Il sorprendente risultato è che gli sperimentatori sono in grado di prevedere, con un anticipo di 7-10 sec., se la decisione dei soggetti sperimentali sarà quella di premere il bottone destro o il sinistro. Apparentemente, un notevole problema per i fautori del libero arbitrio: si pensi come sarebbe sorprendente se, mentre siamo seduti al ristorante e pensiamo cosa ordinare, il cameriere sappia con 7-10 secondi di anticipo cosa sceglieremo!

Questo esperimento è molto interessante e in effetti rappresenta un'evoluzione rispetto agli esperimenti di Libet sia perché dal punto di vista tecnologico è più avanzato (utilizza la risonanza magnetica funzionale invece dell'elettroencefalogramma) sia perché evita un'obiezione importante a cui Libet non riusciva a rispondere. Nell'interpretare i suoi esperimenti, infatti, Libet faceva arbitrariamente iniziare la catena causale rilevante per la decisione nel momento dell'attivazione del potenziale di prontezza nei soggetti, ignorando del tutto la loro precedente decisione consapevole di aderire alla richiesta di prestarsi all'esperimento (e su *quella* decisione gli esperimenti di Libet non dicevano nulla). Soon e colleghi, invece, evitano questa obiezione perché al momento dell'assegnazione del compito sperimentale i soggetti devono limitarsi ad accettare di compiere la scelta in un momento successivo ossia quando il soggetto sarà monitorato sperimentalmente. In generale, secondo molti autori, gli esperimenti di Libet mostrano che la coscienza si limita a ratificare una decisione già determinata dall'attività cerebrale inconscia (su cui noi non possiamo ovviamente esercitare alcun controllo razionale). E ciò proverebbe che le decisioni libere non esistono e che, dunque, il libero arbitrio è una mera illusione.

Varie ragioni, tuttavia, inducono a pensare che esperimenti come quelli di Libet e di Soon, per quanto interessanti e degni di analisi, non hanno conseguenze molto rilevanti rispetto alla questione del libero arbitrio (mentre sono rilevanti rispetto all'indagine sulla fenomenologia degli stati coscienti). In primo luogo, infatti, i sostenitori della veneranda tradizione filosofica detta "compatibilismo" – la classica concezione sviluppata da Locke, Hume e Leibniz e oggi difesa, tra gli altri, da Daniel Dennett, secondo cui il determinismo è conciliabile con il libero arbitrio – non

hanno alcuna difficoltà a concedere che le nostre decisioni siano totalmente determinate da fattori al di là del nostro controllo: questo, infatti, è proprio ciò che il determinismo implica. Almeno una concezione classica del libero arbitrio, dunque, resterebbe in piedi anche se la tesi principale di questi esperimenti fosse dimostrata. Diverso naturalmente sarebbe il caso se questi esperimenti avessero dimostrato che la nostra volontà cosciente, oltre a essere determinata da fattori inconsci, non ha alcuna efficacia causale nella generazione dell'azione in questione (cioè se mostrasse che la mente è *epifenomenica*). In questo esperimento, però, la volontà cosciente – sebbene causalmente determinata ed entri in gioco in ritardo rispetto ai fattori causali inconsci – sembra giocare un ruolo decisivo nella causazione dell'azione: senza la decisione cosciente, infatti, l'azione non verrebbe compiuta.

Contro la tesi che gli esperimenti di Libet e Soon dimostrino l'illusorietà del libero arbitrio si possono inoltre muovere anche altre critiche. Come già notato da Bennett e Hacker¹³, per esempio, questo tipo di esperimenti si incentrano sul momento in cui nel soggetto insorge la consapevolezza dell'impulso a piegare il proprio dito o a premere il bottone: il momento della decisione cosciente, insomma, è identificato con quello in cui il soggetto avverte l'impulso a compiere l'azione richiesta. In realtà, però, l'impulso a compiere un'azione non è né condizione necessaria né condizione sufficiente per il compimento di un'azione che noi considereremmo libera. Non è condizione necessaria (e dunque possono esserci azioni volontarie senza che il soggetto avverta l'impulso a compierle) perché molto spesso, quando compiamo un'azione che noi giudichiamo libera, non avvertiamo alcun impulso a compierla: si pensi a quando, guidando, sterziamo per curvare o a quando, mangiando portiamo una posata verso la bocca o, ancora, a quando pronunciamo intenzionalmente una frase durante una normale conversazione. D'altra parte, la presenza dell'impulso ad agire non è nemmeno sufficiente perché segua un'azione che consideriamo libera: spesso, infatti, un tale impulso precede azioni non volontarie, come quando ci viene da starnutire in pubblico o quando sbadigliamo di fronte a un interlocutore poco brillante.

Ma c'è di più. Non solo il tipo di situazioni sperimentali studiate da Libet e Soon non è tipico di una decisione *libera*, ma non lo è nemmeno di una *decisione* genuina. Il problema è che non è giustificato richiamare

¹³ M. Bennett, P. Hacker, *Philosophical Foundations of Neuroscience*, Blackwell, Oxford, 2003; M. De Caro, *Is Emergentism Refuted by the Neurosciences? The Case of Free Will*, in A. Corradini, T. O'Connor (eds.), *Emergence in Science and Philosophy*, Routledge, London, 2010, pp.190-211; M. De Caro, M. Marraffa, *Mente e morale*, Luiss University Press, Roma 2016, cap. 1.

il concetto di decisione quando l'alternativa posta ai soggetti sperimentali si riduce a scelte banali e irrilevanti. Come postula la teoria delle decisioni (che su questo, peraltro, si muove in accordo con il senso comune), la nozione di decisione è infatti concettualmente correlata a quella di preferenza: ogni decisione presuppone una gerarchia di preferenze, in base alla quale il soggetto valuta i vari corsi d'azione che gli si presentano. Se al ristorante devo scegliere se prendere frutta o insalata, valuto i pro e i contro di queste due opzioni, e giudico in base al risultato di questa valutazione: ed è proprio per questo che sono conscio di aver preso una decisione. Se invece al supermercato devo scegliere se prendere il carrello di sinistra o quello di destra, la cosa mi è del tutto indifferente e agisco con un automatismo (se mi chiedessero perché ho preso un carrello invece che l'altro, risponderai che non lo so, oppure che la cosa era del tutto irrilevante e ho scelto a caso). Ora, le presunte "decisioni" cui fanno riferimento Libet e Soon – piegare il dito indice oppure premere il pulsante destro o quello sinistro – non hanno alcun interesse per i soggetti sperimentali: essi non hanno preferenze, in un senso o nell'altro, che possano essere poste in gerarchia per operare la scelta. Pertanto non si tratta di scelte genuine – e dunque non sono rilevanti per la questione del libero arbitrio.

Va sottolineato, infine, che l'esperimento di Soon si limita a mostrare che un determinato *pattern* neurologico predice con un 60% di accuratezza se la decisione che i soggetti prenderanno sarà quella di premere il bottone destro o il sinistro. Questo risultato è evidentemente significativo dal punto di vista statistico (perché è ben sopra il 50%); tuttavia, con quest'ordine di grandezza esso non sarebbe di per sé in grado di compromettere l'idea di libero arbitrio persino se i risultati dell'esperimento si potessero applicare a *tutte* le decisioni umane (il che, naturalmente, è tutto da dimostrare). Non vi è infatti alcuna ragione di credere – né mai nessun filosofo lo ha fatto – che gli esseri umani siano integralmente liberi né che tutte le scelte che essi compiono siano libere; e al massimo l'esperimento mostra che nel 60% dei casi i soggetti sono determinati nelle loro decisioni. Resta dunque un 40% di indeterminazione: una percentuale della quale chi volesse continuare a difendere il cosiddetto "libertarismo" (la concezione, già di Epicuro, Kant e oggi di Searle, secondo cui il libero arbitrio presuppone l'indeterminismo) potrebbe tranquillamente accontentarsi.

Gli esperimenti della tradizione di Libet, in definitiva, non sembrano portare sostegno ai nemici del libero arbitrio. Anche ignorando tutte le obiezioni viste sopra, infatti, essi possono tutt'al più mostrare che le nostre scelte sono determinate: ma questo, come abbiamo visto, non è un problema per i fautori del libero arbitrio che simpatizzano per il compatibilismo.

Se dunque la sfida contro il libero arbitrio (e, di conseguenza, contro le attribuzioni di responsabilità e contro la teoria retributivistica della pena che tali attribuzioni presuppone) si fondasse solo su questo tipo di esperimenti, essa non parrebbe destinata al successo.

3. *La sfida dell'epifenomenismo*

Più incisiva è la sfida che alla galassia di concetti su cui si fonda il retributivismo è posta da un diverso tipo di esperimenti, che hanno portato alla luce la discrepanza fra i *motivi* (esplicativi) che il soggetto adduce per spiegare le proprie azioni e i *fattori motivazionali* inconsapevoli (ossia le cause reali) delle azioni stesse¹⁴. A differenza degli studi di Libet e Soon, questo tipo di studi suggeriscono che vi sono condizioni in cui, al di là delle apparenze, gli stati coscienti non sono causalmente rilevanti nella generazione delle azioni e dunque che, almeno in alcuni casi, la nostra mente è epifenomenica; e, come detto sopra, la minaccia dell'epifenomenismo è assai rilevante per la questione del libero arbitrio.

Alcuni autori sostengono che in realtà la mente è *sempre* epifenomenica, ossia i nostri stati mentali coscienti non determinano mai alcunché (né decisioni né azioni, dunque). Lo psicologo statunitense Daniel Wegner¹⁵ è stato uno dei maggiori rappresentanti di questo indirizzo di ricerca. A suo giudizio, l'esperienza fenomenologica della volontà cosciente prende corpo in seguito a un'inferenza "retrodittiva": dopo aver compiuto un'azione, o mentre la sta compiendo, il soggetto inferisce che la responsabilità causale dell'azione vada attribuita a un determinato "pensiero", per esempio, una decisione – e ciò a prescindere dalla correttezza dell'inferenza.

Uno dei principali esperimenti su cui si basa la concezione epifenomenistica della mente è stato condotto nel 2005 da Petter Johansson *et al.*¹⁶ dell'Università di Lund. Questo gruppo propone un nuovo paradigma

¹⁴ Questa sezione riprende, semplificandola, la discussione svolta in M. De Caro, M. Marraffa, *Mente e morale*, cit.

¹⁵ D. Wegner, *The Illusion of Conscious Will*, Harvard University Press, Cambridge (MA), 2002.

¹⁶ P. Johansson, L. Hall, S. Sikström, A. Olsson, *Failure to detect mismatches between intention and outcome in a simple decision task*, in «Science», 310, 2005, pp. 116-119; cfr. anche P. Johansson, L. Hall, B. Tärning, S. Sikström, N. Chater, *Choice Blindness and Preference Change: You Will Like This Paper Better If You (Believe You) Chose to Read It!*, in «Journal of Behavioral Decision Making».

teorico, quello “*Choice blindness*”. Nell’esperimento di Johansson, ai soggetti sperimentali sono presentate le foto di due persone dello stesso sesso e viene loro chiesto di indicare quale giudicano la più interessante; una volta compiuta la scelta, i partecipanti sono invitati a dichiarare le ragioni della loro preferenza. Alcune prove sono però manipolate: lo sperimentatore, trasformandosi in prestigiatore, scambia le due immagini e chiede ai soggetti, ignari del trucco, di esplicitare le ragioni della loro preferenza in relazione alla foto che *non* hanno scelto. Sorprendentemente, solo nel 28% delle prove manipolate, i partecipanti si sono resi conto dello scambio; nel restante 72% delle prove i soggetti, ignari dello scambio di foto, hanno invece offerto spiegazioni volte a giustificare una scelta che non avevano mai compiuto, spesso menzionando fattori che non potevano far parte delle ragioni per la scelta (per esempio, un partecipante spiegò di aver scelto il volto che gli era stato mostrato dicendo «mi piacciono i suoi orecchini», quando in realtà la donna nell’immagine scelta non indossava orecchini). I resoconti introspettivi nelle due condizioni, si noti, non hanno evidenziato alcuna differenza rispetto al grado di coinvolgimento emotivo, al livello di dettaglio della descrizione e al grado di sicurezza che i partecipanti hanno espresso rispetto alla loro scelta. Inoltre, dal confronto fra le risposte dei soggetti nella condizione di scelta effettiva e in quella manipolata non è emersa alcuna differenza sistematica. Questo, dunque, è un caso inequivocabile di confabulazione delle motivazioni di una scelta: ossia un caso in cui la mente è epifenomenica.

Secondo Wegner e Johansson, questo genere di studi suggerisce che noi siamo sempre ignari dei motivi per cui compiamo le scelte che compiamo. Le ragioni che offriamo in favore delle nostre scelte, insomma, sarebbero sempre congegnate a posteriori, per dare conto di quanto accaduto come se l’avessimo veramente deciso noi. Seconda questa prospettiva, le confabulazioni servono in primo luogo al soggetto agente, che ha bisogno di confermarsi nella sua fiducia intuitiva di essere in controllo delle scelte e delle azioni che compie.

Se le teorie di Wegner e di Johansson si dimostrassero corrette, la tesi del libero arbitrio sarebbe definitivamente confutata. Si noti: la concezione compatibilistica citata sopra (quella che dà conto del libero arbitrio in un contesto deterministico) si fonda sul fatto che le nostre scelte coscienti, pur essendo interamente determinate, esercitano un potere causale sulle nostre decisioni e, dunque anche sulle nostre azioni. Nel quadro esplicativo proposto da Wegner e Johansson, invece, nessuna nostra decisione cosciente ha mai alcun potere causale. In quel quadro, dunque, noi non siamo liberi e, conseguentemente, nemmeno responsabili. Se questi avessero ragione, i

fondamenti concettuali del retributivismo crollerebbero: e dunque quanti perorano l'adozione di una concezione strettamente utilitaristica della pena avrebbero vinto la loro battaglia.

È importante dunque chiedersi se la concezione epifenomenistica sia corretta. In realtà, non è difficile notare che Wegner e Johansson estrapolano dai loro dati un'interpretazione universale che quei dati non sono in grado di sostenere. La loro tesi è che l'esperienza della volontà cosciente – per mezzo della quale noi ci consideriamo autori delle nostre azioni – sia *sempre* ingannevole, nel senso che noi ci illudiamo di causare le azioni che compiamo, mentre l'efficacia causale è propria esclusivamente dei processi mentali che per loro natura sono, e sempre rimarranno, inconsci. Tuttavia gli esperimenti da loro presentati, ancorché sorprendenti, riguardano ambiti molto limitati dell'agire umano: la loro conclusione, dunque, si fonda su un'audace estrapolazione dal particolare al generale. Inoltre, in buona parte, i casi studiati proposti questi esperimenti riguardano situazioni artificiali, in cui i processi di deliberazione non concernono cose che ci stanno veramente a cuore e su cui dunque pondereremmo attentamente. Infine Wegner e Johansson presuppongono una tesi fenomenologicamente dubbia, poiché attribuiscono al senso comune (e alla filosofia) la tesi che le azioni volontarie siano causate dagli stati coscienti che abbiamo nel momento in cui pensiamo di causarle. In realtà, una cosa è la causazione volontaria delle azioni, una cosa diversa la nostra consapevolezza di tale causazione: e non è affatto detto che le due coincidano. Anche l'attacco al retributivismo basato sulla epifenomenistica della mente manca dunque il suo bersaglio¹⁷.

4. La teoria della pena tra utilitarismo e retributivismo

Né la sfida deterministica né quella epifenomenistica contro le idee di libertà e di responsabilità (e conseguentemente contro il retributivismo) hanno successo. Tra le due concezioni, tuttavia, va ribadita un'importante asimmetria. Se il determinismo fosse vero, ai difensori della libertà (e del retributivismo) rimarrebbe a disposizione la concezione compatibilistica

¹⁷ Per approfondire questi temi, cfr. A. Mele, *Free: Why Science Hasn't Disproved Free Will*, Oxford University Press, New York, 2014, trad. it. *Liberi. Perché la scienza non ha confutato il libero arbitrio*, Carocci, Roma, 2015. M. De Caro, A. Lavazza, *Free Will as an Illusion: Ethical and Epistemological Consequences of an Alleged Revolutionary Truth*, in «Social Epistemology Review and Reply Collective», 3, 2014, pp. 40-50.

della libertà che, appunto, non vede nel determinismo un ostacolo per il libero arbitrio, a condizione che gli stati mentali coscienti incidano causalmente sul mondo. Se fosse vero l'epifenomenismo, invece, gli stati mentali consci non avrebbero nessuna efficacia causale: pertanto il compatibilismo sarebbe inapplicabile e, di conseguenza, libero arbitrio e responsabilità si dimostrerebbero del tutto illusori. E parimenti cadrebbe il retributivismo, sia positivo sia negativo: i difensori di questa concezione, più che del determinismo, devono dunque preoccuparsi dell'epifenomenismo, sperando che in futuro non si raccolgano dati adeguati a ratificarne la correttezza. Data la possibilità, sia pure al momento remota, che un giorno l'epifenomenismo si dimostri corretto, e dunque il retributivismo debba essere abbandonata, resta dunque aperta una questione: è vero, come abbiamo visto sostenere da Cohen e Greene e dai loro colleghi, che se venisse basato unicamente su elementi utilitaristici il sistema penale guadagnerebbe in umanità?

In realtà, ci sono ottime ragioni per pensare che la risposta a questa domanda debba essere negativa. In primo luogo, va notato come Cohen, Greene e colleghi idealizzino eccessivamente il modo con cui la società tratta chi ha commesso crimini senza avere la capacità di intendere e di volere. In secondo luogo, e questo è il punto più interessante in questa sede, questi autori sono ingiustificatamente ottimisti rispetto alle conseguenze di una concezione radicalmente utilitaristica della pena. La questione cruciale è che, con lo strumentario concettuale dell'utilitarismo, non si riesce a dimostrare che – quali che siano le circostanze – è *sempre* ingiusto punire un innocente o punire spropositatamente un colpevole. Vediamo perché.

La forma più tradizionale di utilitarismo è "utilitarismo dell'atto", la concezione secondo cui, per comportarsi moralmente, si devono compiere le azioni che massimizzano l'utilità generale (che è generalmente interpretate come felicità generale). Questa definizione, che pure ha il vantaggio di essere molto semplice, lascia sfortunatamente spazio per le pratiche intuitivamente ingiuste come quelle del capro espiatorio e dell'irrogazione di pene spropositate: e, in proposito, basterà citare il caso in cui qualcuno è punito anche se è innocente o è punito in modo spropositato se è colpevole, e ciò in funzione di deterrenza per altri potenziali criminali. Si possono immaginare facilmente situazioni in cui l'irrogazione di queste pene – intuitivamente ingiuste – rappresenterebbe un beneficio generale per la comunità e in cui, dunque, esse sarebbero *ipso facto* accettabili nella prospettiva dell'utilitarismo dell'atto. Il che mostra che la mera massimizzazione dell'utilità sociale non può rappresentare l'unico *standard* di correttezza della teoria della pena, come invece sostiene

l'utilitarismo dell'atto.

Per questo, alcuni autori hanno sviluppato una versione alternativa dell'utilitarismo, l'"utilitarismo della regola", secondo cui per comportarsi moralmente occorre compiere le azioni che si conformano a regole la cui applicazione garantisca la massimizzazione dell'utilità generale. Tuttavia, per dimostrare che questa concezione è in grado di replicare in modo convincente all'obiezione del capro espiatorio e delle pene spropositate, si dovrebbe dimostrare che una norma che lasciasse la possibilità, in determinate condizioni, di punire un innocente o di punire spropositatamente un colpevole non potrebbe massimizzare l'utilità generale (perché nel caso contrario, alla luce dell'utilitarismo della regola, tale regola dovrebbe essere comunque adottata). A prima vista, in effetti, può sembrare che le cose stiano proprio così: se i potenziali criminali sapessero che potrebbero essere puniti anche nel caso in cui non commettessero alcun crimine, non verrebbero certo scoraggiati dal commetterne. Inoltre, si potrebbe sostenere che la punizione di individui innocenti potrebbe generare un'indignazione generalizzata da parte della comunità; e certamente un tale indignazione non sarebbe un buon modo di massimizzare l'utilità generale.

Nondimeno, a un'analisi più attenta si può constatare che l'utilitarismo della regola non può veramente risolvere i problemi del capro espiatorio e delle pene spropositate. Si possono infatti immaginare casi in cui l'utilità generale sarebbe massimizzata dall'accettazione di una regola che preveda la possibilità, in condizioni estreme, della punizione di un innocente o dell'irrogazione di una pena sproporzionata. In questo senso si può per esempio menzionare il caso storico della decimazione, che fu pratica abbastanza comune nella Prima guerra mondiale negli eserciti dell'Intesa e in quello italiano, e niente affatto comune, invece, negli eserciti delle potenze centrali¹⁸. La pratica della decimazione consisteva nell'esecuzione di alcuni soldati scelti a caso da una compagnia, in genere perché, nel suo complesso, quella compagnia aveva dimostrato scarso entusiasmo bellico verso il nemico. Lo scopo di questa pena era di fornire un monito indelebile ai compagni dei soldati giustiziati. In proposito va notato che, data la procedura di selezione casuale dei soldati da giustiziare, poteva certamente capitare che venissero fucilati soldati che certamente non si erano dimostrati codardi (come ben mostrato in *Orizzonti di gloria*, il capolavoro di Stanley Kubrick, basato su una storia vera accaduta sul fronte francese). Tuttavia si potrebbe argomentare, in modo niente

¹⁸ I. Guerrini, M. Pluviano, *Le fucilazioni sommarie nella Prima guerra mondiale*, Gaspari, Udine, 2004.

affatto implausibile, che la pratica della decimazione può aver contribuito alla vittoria di paesi alleati e, dunque, che essa possa aver massimizzato l'utilità collettiva in quei paesi. Ma ciò dimostrerebbe forse che la pratica della decimazione sia moralmente accettabile¹⁹? La risposta è, senza alcun dubbio, negativa.²⁰

Riassumendo: il problema principale di ogni forma di utilitarismo è che l'ideale di giustizia non può essere tradotto senza residui nei termini dell'utilità generale. Pertanto, se come fondamento della teoria della giustizia si accettasse una forma pura di utilitarismo, ci correrebbe sempre il rischio di giustificare pratiche intuitivamente immorali come quelle del capro espiatorio e delle pene spropositate.

Per fortuna abbiamo a disposizione un'alternativa molto più promettente: quella di limitare l'utilitarismo dell'atto per mezzo della retribuzione negativa. In questa prospettiva, per esempio, si sono mossi, due giganti del pensiero etico-giuridico anglosassone del Novecento, John Rawls²¹ e H. L. A. Hart²². Hart, in particolare, ha offerto la trattazione più convincente della questione. A suo giudizio, la giustificazione della pena può avvenire solo su un fondamento utilitaristico: si possono punire soltanto coloro che è utile punire (una tesi, si noti, ovviamente incompatibile con il retributivismo positivo, secondo cui si devono punire tutti coloro che lo meritano, quali che siano le conseguenze della loro punizione). Per quanto riguarda, però, la distribuzione delle pene da parte dei giudici, Hart pone invece un vincolo retributivistico-negativo. In questa prospettiva, stante l'utilità della pena, non si deve mai punire chi non lo merita.

Il punto cruciale, pertanto, è che, se le idee di libero arbitrio, responsabilità morale e merito si dimostrassero illusorie, allora, oltre al retributivismo positivo, cadrebbe anche il retributivismo negativo (secondo cui non si può punire nessuno che non lo meriti e in una misura superiore a quanto meriti). E in questo modo non ci sarebbero più argini, come

¹⁹ Si potrebbe forse obiettare che questo ragionamento è ipotetico, perché non è certo che la pratica della decimazione contribuì alla vittoria alleata. Naturalmente, però, in questa sede la questione è di principio: serve cioè a mostrare che ci possono essere casi in cui la massimizzazione dell'utilità generale non necessariamente produce una pena giusta. In secondo luogo, non sarebbe certo soddisfacente decidere della liceità morale della decimazione a partire della sua utilità empirica!

²⁰ Sulle notevoli difficoltà teoriche cui va incontro l'utilitarismo della regola, cfr. R. Arneson, *Sophisticated Rule Consequentialism: Some Simple Objections*, in «Philosophical Issues», 15, 2005, pp. 235-251.

²¹ J. Rawls, *Two Concepts of Rules*, in «Philosophical Review», 64, 1955, pp. 3-32.

²² H.L.A. Hart, *Punishment and Responsibility*, Oxford University Press, Oxford, 1968.

quelli posti da Hart, alla possibilità che le pratiche del capro espiatorio e delle pene spropositate, inflitte a fini di deterrenza, vengano considerate accettabili: come si è visto sopra, infatti, né l'utilitarismo dell'atto né quello della regola hanno le risorse concettuali per mostrare che tali pratiche sono ingiuste.

Pur assumendo un quadro utilitaristico di sfondo, allora, solo il retributivismo negativo può porre un argine teorico alla possibilità che diventino accettabili pratiche ingiuste, ma utili socialmente, quali la condanna di chi non lo merita o l'inflizione di pene esagerate. Illudersi che le cose stiano diversamente, e addirittura forzare in questa direzione dati scientifici ancora largamente insufficienti, pone il duplice problema di rischiare l'inconsistenza filosofica e l'ingiustizia sociale.

Libero arbitrio e colpevolezza nell'ordinamento giuridico italiano

LORENZO MILAZZO*

SOMMARIO: 1. Una premessa – 2. Il Codice Rocco e i lavori preparatori – 3. Principio di colpevolezza e libero arbitrio nell'ordinamento repubblicano – 4. Una prima difficoltà, d'ordine logico – 5. Dalla libertà del volere alla «motivabilità mediante norme» – 6. Dalla possibilità di assumere l'atteggiamento doveroso della volontà alla constatazione del fatto che *gli altri* in genere lo fanno – 7. 'Come se'.

1. *Una premessa*

Come è noto Daniel C. Dennett è un *compatibilista*¹ e Gregg D. Caruso un *incompatibilista radicale (hard)*: Dennett crede che il libero arbitrio sia compatibile (solo) con il determinismo, Caruso crede che il libero arbitrio sia incompatibile tanto con il determinismo quanto con l'indeterminismo². Avendo la pazienza di leggere fino in fondo il libro che hanno scritto insieme su libero arbitrio e responsabilità morale³ (ne serve parecchia, ma vale la pena averla) probabilmente ci si farà l'idea che la disputa fra i due e forse, più in generale, la stessa controversia fra compatibilisti e incompatibilisti, abbia natura normativa e non teoretica e verta sulle condizioni alle quali sia giusto ascrivere a ciascuno la responsabilità dei propri atti, lodarlo o biasimarlo e, talora, anche punirlo.

Chi non abbia, d'altra parte, la pretesa di stabilire se e quando sia giusto punire e riprovare né, a maggior ragione, quella di convincere ogni altro della bontà delle proprie convinzioni, ma creda piuttosto di doversi accontentare di provare a capire a quali condizioni questo o quell'ordinamento giuridico prescriva ai propri organi di infliggere una pena a chi violi certe sue

* Dipartimento di Giurisprudenza, Università di Pisa.

¹ Cfr., ad esempio, D.C. Dennett, *Freedom Evolves*, Viking Press, New York, 2003, trad. it. di M. Pagani, *L'evoluzione della libertà*, Cortina, Milano, 2004, p. 132.

² Cfr., ad esempio, G.D. Caruso, *Rejecting Retributivism. Free Will, Punishment, and Criminal Justice*, Cambridge University Press, Cambridge-New York, 2021, pp. 15 e 70.

³ D.C. Dennett, G.D. Caruso, *Just Deserts. Debating Free Will*, Polity Press, Cambridge, 2021, trad. it. di S. Colloca, *A ognuno quel che si merita. Sul libero arbitrio*, Raffaello Cortina, Milano, 2022.

norme (o li autorizzi a farlo), potrà forse evitare di passare da una «pompa delle intuizioni» all'altra per mostrare ai suoi interlocutori che in realtà non pensano quello che pensano di pensare e potrà richiamare piuttosto l'attenzione su alcuni testi dai quali generalmente si ritiene possa essere desunto, per così dire, il punto di vista di ciascun ordinamento (che non è poi che quello di coloro che di volta in volta lo impersonano, siano essi giudici, giuristi, legislatori, ecc.). Le conclusioni alle quali perverrà saranno comunque incerte e controverse e continueranno a dipendere in larga parte dalle sue precomprensioni e dagli assunti normativi da cui muove, ma potranno forse procurare qualche guadagno teorico che, per quanto esiguo, sarebbe di per sé precluso a chi continuasse a coltivare la speranza di decidere una volta per tutte quale genere di libertà serva a fondare la responsabilità morale e a giustificare la riprovazione, il biasimo e la pena.

I criteri di imputazione della responsabilità penale variano ovviamente «a seconda del momento storico, delle condizioni politiche, delle forme di Stato e delle specificità culturali di ciascuno Paese»⁴, e non di rado gli interpreti esprimono opinioni discordanti riguardo agli stessi ordinamenti: così ad esempio, se Stephen J. Morse, che insegna diritto e psicologia all'Università della Pennsylvania, ritiene, come è noto, che «solo il compatibilismo [possa] spiegare e giustificare le nostre dottrine e le nostre pratiche morali e giuridiche»⁵ e che «libero arbitrio e causalità non [siano] criteri di responsabilità o di giustificazione»⁶, Luis E. Chiesa, dell'Università di Buffalo, è dell'idea che il diritto penale statunitense presupponga «l'esistenza di attori dotati di libero arbitrio e capaci di controllare significativamente la propria condotta» e che perciò «il libero arbitrio sia incorporato» nella sua stessa «struttura»⁷.

⁴ M. Donini, L. Ramponi, *Il principio di colpevolezza*, in G. Insolera, N. Mazzacava e M. Pavarini (a cura di), *Introduzione al sistema penale*, Giappichelli, Torino, 2012, p. 285.

⁵ S.J. Morse, *New Neuroscience, Old Problems*, in B. Garland (ed.), *Neuroscience and the Law. Brain, Mind, and the Scales of Justice*, Dana Press, New York, 2004, p. 173. Cfr. anche, *ivi*, p. 176: «Compatibilism [...] offers the most plausible causal metaphysics and the only coherent positive explanation of our current legal (and moral) practices». E alla pagina seguente: «...the presence or absence of contra-causal freedom or something like this is not part of the criteria of any legal doctrine that holds some people nonresponsible».

⁶ *Ivi*, p. 179.

⁷ L.E. Chiesa, *Punishing Without Free Will*, in «Utah Law Review», 4, 2011, p. 1406. A conclusioni simili erano pervenuti qualche anno prima J. Greene e J. Cohen in *For the Law, Neuroscience Changes Nothing and Everything*, in «Philosophical Transactions of the Royal Society B: Biological Sciences», 359, 2004, nonché in S. Zeki, O. Goodenough (eds.), *Law and the Brain*, Oxford University Press, New York, 2006. Cfr. in particolare, *ivi*, p. 209 e A. Farano, *Naturalizzare la responsabilità. La teoria della responsabilità alla prova delle neu-*

Benché le posizioni di Morse siano molto note in Italia, oltreoceano godono forse di minor fortuna. Sam Harris ricorda che nel 1978, in *United States v. Grayson*, la Corte suprema ravvisò nel «libero arbitrio un fondamento “universale e persistente”» del sistema giuridico statunitense, chiarendo che il determinismo è incompatibile con i suoi principi basilari⁸. Anche Caruso lo rileva, constatando che «il diritto penale è fondato sull'idea che le persone possano essere considerate moralmente responsabili per le loro azioni perché hanno scelto liberamente di compierle»⁹ e che tanto negli Stati Uniti quanto, ad esempio, nel Regno Unito, in Australia, in Canada, nella Nuova Zelanda e in Israele, la pena ha ancora fondamento *retributivo* e la sua infrazione continua ad essere giustificata dalla «blameworthiness of offenders»¹⁰.

2. Il Codice Rocco e i lavori preparatori

Un rapido sguardo al Codice Rocco e ai suoi lavori preparatori potrebbe indurre a credere che il nostro ordinamento muova da presupposti più o meno esplicitamente compatibilisti.

Nella *Relazione a S.M. il Re del Ministro Guardasigilli presentata nell'udienza del 19 ottobre 1930 per l'approvazione del testo definitivo del codice penale* Alfredo Rocco, dando conto dei problemi di coordinamento fra gli art. 42 e 85 emersi in Commissione, informa Vittorio Emanuele III che la proposta «di qualificare la volontà come libera» non era stata accolta perché avrebbe potuto «far dubitare che col termine “libera” si [volesse] alludere al libero arbitrio, questione filosofica insoluta e forse insolubile, che deve rimanere del tutto estranea al diritto penale, in quanto la soluzione di essa non è affatto necessaria per la giustificazione razionale e per l'applicazione della legge penale»¹¹. Poco prima i termini della

rosienze, in R. Brighi, S. Zullo (a cura di), *Filosofia del diritto e nuove tecnologie. Prospettive di ricerca tra teoria e pratica*, Aracne, Roma, 2015, p. 171.

⁸ S. Harris, *The Moral Landscape: How Science Can Determine Human Values*, Free Press, New York-London-Toronto-Sidney, 2010, p. 106.

⁹ G.D. Caruso, *Rejecting Retributivism*, cit., p. 3.

¹⁰ Ivi, pp. 3 e 7. Per Massimo Donini, del resto, «in Inghilterra, ma anche oltre oceano, overseas, che il diritto penale, quale istituzione “of blame and punishment”, sia una morale applicata, appare evidente» (M. Donini, *An Impossible Exchange? Prove di dialogo tra civil e common lawyers su legalità, morale e teoria del reato*, in «Rivista italiana di diritto e procedura penale», LXIII, 2017, p. 25; cfr. anche p. 31).

¹¹ In «Gazzetta ufficiale», n. 251 del 26 ottobre 1930, p. 4459.

questione sono illustrati tanto chiaramente da non lasciare, in apparenza, alcun margine di dubbio.

«Se il reato non può concepirsi se non sul fondamento della coscienza e volontarietà dell'azione od omissione umana [...], appare evidente che autori di reato non possono considerarsi dalla legge se non coloro che hanno la capacità di intendere o di comprendere o di discernere e la capacità di coscientemente volere»¹². Perciò, prosegue Rocco, «la responsabilità penale [...] rimane, anche nel nuovo codice, saldamente affidata al principio dell'imputabilità psichica e morale dell'uomo, fondato a sua volta sulla normale capacità, e quindi sulla libertà d'intendere e di volere»¹³.

Questo non significa tuttavia, come subito il Guardasigilli si affretta a precisare, che si sia voluto in questo modo riprodurre «il vecchio e vieto concetto metempirico e scolastico del così detto *arbitrium indifferentiae*, secondo il quale la volontà umana si sottrarrebbe, essa sola, fra tutti i fenomeni del mondo empirico, all'impero della legge universale di causalità e il volere umano dovrebbe concepirsi come un effetto senza causa che si determina in assenza di qualsiasi motivo»¹⁴. Gli estensori del codice non dubitano che «il volere dell'uomo subisc[a] [...] la legge di causalità al pari di ogni altro fenomeno»¹⁵. Solo che, diversamente da quel che era stato sostenuto da alcuni fra i più autorevoli esponenti della scuola positiva, «tale causalità non è fisica o materiale, e neppure fisiologica, [...] ma meramente psicologica»¹⁶. Perciò, per quanto la volontà umana sia «determinata, [...] questa sua determinazione non è che un'autodeterminazione, cioè una facoltà di determinarsi per motivi consci in conformità del carattere individuale di ciascuno»¹⁷.

Il codice e i suoi estensori temono il «determinismo [...] fisico o meccanico o fisiologico»¹⁸ ma non quello «psicologico»¹⁹, in cui ravvisano

¹² Ivi, p. 4449.

¹³ *Ibid.*

¹⁴ *Ibid.*

¹⁵ *Ibid.*

¹⁶ *Ibid.*

¹⁷ *Ibid.* Cfr. A. Manna, *L'imputabilità e i nuovi modelli di sanzione. Dalle "finzioni giuridiche" alla "terapia sociale"*, Giappichelli, Torino, 1997, p. 33.

¹⁸ *Relazione a S.M. il Re del Ministro Guardasigilli*, cit., p. 4449. Sulla «minaccia dell'epifenomenismo» si può vedere M. De Caro, *Colpevolezza e scienze cognitive*, in «Ragion pratica», 2, 2020, pp. 75 e ss. e, in questo volume, *Retributivismo, utilitarismo e libero arbitrio*, pp. 59 e ss.

¹⁹ *Relazione a S.M. il Re del Ministro Guardasigilli*, cit., p. 4449.

piuttosto un presupposto necessario del funzionamento del diritto²⁰. Nel *Discorso al Senato del Regno* Rocco non avrebbe potuto essere più chiaro: «Tutte le sanzioni giuridiche, [...] e quindi anche le sanzioni penali, agiscono ponendo motivi alla volontà, in modo da *determinarla* in uno piuttosto che in un altro senso. [...] Tutto il diritto, e quindi anche il diritto penale, presuppone l'autodeterminazione per motivi consci, il determinismo psicologico»²¹.

Filippo Grisogni lo rilevò nel suo *Corso di diritto penale secondo il nuovo codice*²², ma qualche tempo dopo Luigi Scarano e Biagio Petrocelli tornarono sulla questione proponendo una curiosa interpretazione d'ispirazione 'incompatibilista' e 'libertaria' del codice e delle relazioni del Guardasigilli²³.

Secondo Petrocelli, in particolare, per comprendere appieno le posizioni espresse da Alfredo Rocco nella sua *Relazione al Re* riguardo al libero arbitrio sarebbe stato necessario andare a rivedere gli atti della Commissione ministeriale incaricata di dare un parere sul Progetto preliminare, e in particolare le dichiarazioni rese nella seduta dell'11 marzo 1928 da Arturo Rocco, le quali dimostravano a suo avviso inconfutabilmente che ponendo «il concetto di determinismo psichico quale fondamento dell'indirizzo del Progetto» si intendeva in realtà basare anche il nuovo codice sull'assunto della libertà del volere²⁴. Già Scarano, del resto, tenendo conto delle

²⁰ Cfr. A. Rocco, *Sulla delega al governo del Re della facoltà di emendare il codice penale, il codice di procedura penale, le leggi sull'ordinamento giudiziario e di apportare nuove modificazioni e aggiunte al codice civile. Discorso al Senato del Regno*, in Id., *La trasformazione dello Stato. Dallo Stato liberale allo Stato Fascista*, La Voce, Roma, 1927, p. 292: «[...] ammettiamo quel determinismo, che è condizione necessaria e sufficiente per concepire la pena come una sanzione giuridica, vale a dire il determinismo psichico». E alla pagina seguente: «[...] esiste una volontà, che si determina immediatamente per motivi consci. Affermando l'esistenza di una tale volontà, si afferma la possibilità d'influire su di essa, di creare una *causa psicologica*, un motivo, capace di *determinarla*» (corsivi aggiunti).

²¹ *Ibid.* (corsivo aggiunto). Cfr. H. Kelsen, *Reine Rechtslehre*, Verlag Franz Deuticke, Wien, 1960, trad. it. di M. Losano, *La dottrina pura del diritto*, Einaudi, Torino, 1973 (3ª ed.), p. 113: «L'opinione che soltanto la libertà dell'uomo – e cioè il fatto di non essere soggetto alla legge causale – renda possibile la responsabilità e quindi l'imputazione è in aperto contrasto con i dati della vita sociale. Lo statuire un ordinamento normativo che regoli il comportamento degli uomini, in base al quale soltanto è possibile l'imputazione, presuppone addirittura che la volontà degli uomini, di cui si regola il comportamento, sia causalmente determinabile, cioè non libera».

²² F. Grisogni, *Corso di diritto penale secondo il nuovo codice. I. Introduzione*, Cedam, Padova, 1932, p. 105.

²³ L. Scarano, *Libera volontà e libero arbitrio nel diritto penale*, Giuffrè, Milano, 1937 e B. Petrocelli, *La colpevolezza. Lezioni introduttive*, Armanni, Napoli, 1948.

²⁴ B. Petrocelli, *La colpevolezza*, cit., pp. 97-98.

dichiarazioni rese da Arturo quella domenica mattina, era giunto più o meno alle stesse conclusioni²⁵.

Le dichiarazioni di Arturo Rocco a cui si riferiscono Scarano e Petrocelli, tuttavia, non gettano nuova luce sul testo della *Relazione al Re*. Arturo certo affermava nel 1928 che «la volontà si deve concepire come libera» e che «non si può disconoscere il dato psicologico della volontà umana» come facevano, a suo avviso erroneamente, i fautori del «determinismo [...] fisico, o meccanico, o fisiologico»²⁶, ma escludeva anche, proprio come avrebbe fatto Alfredo due anni dopo, che si potesse

concepire una volontà senza causa, una volontà senza motivi, una volontà come un *fiat* che nasce dal nulla, una volontà come mero *arbitrium indifferentiae*. La volontà umana non si sottrae, neppure essa, alla legge di causalità, che governa tutti i fenomeni. Ma c'è differenza fra causalità e causalità, fra determinazione e determinazione. C'è un determinismo fisico o meccanico, che governa i fenomeni fisici o meccanici; c'è un determinismo fisiologico, che governa i movimenti fisiologici del corpo, agendo come stimolo; e poi c'è il determinismo psicologico, che è determinazione secondo cause psicologiche, cioè motivi coscienti, che determinano la volontà umana²⁷.

Secondo Arturo Rocco «la volontà non è puramente ricettizia e passiva di fronte ai motivi, ma reagisce in conformità del carattere, con la possibilità di contrapporre un motivo all'altro e quindi di scegliere fra i vari motivi»²⁸, ma questo non significa che, come vorrebbe Petrocelli, per Rocco la volontà «frontegg[i] e domin[i] l'azione» dei motivi contrastanti che agiscono su di essa «risolvendosi per l'uno o per l'altro mercé l'esercizio di un autonomo potere di scelta, cioè della facoltà di dirigersi in un senso o nell'altro»²⁹ – senza che la sua scelta, dobbiamo immaginare, sia determinata da motivo alcuno? –, perché una volontà siffatta sarebbe precisamente quell'*arbitrium indifferentiae* del quale Arturo, come Alfredo, si voleva sbarazzare.

²⁵ Cfr. L. Scarano, *Libera volontà e libero arbitrio nel diritto penale*, cit., pp. 135-136 e pp. 160-161.

²⁶ *Lavori preparatori del codice penale e del codice di procedura penale. Volume IV. Atti della Commissione ministeriale incaricata di dare parere sul Progetto preliminare di un nuovo codice penale. Parte 11^a. Verbali delle sedute della Commissione (Libro I del Progetto)*, Tipografia delle Mantellate, Roma, 1929, p. 137.

²⁷ *Ibid.*

²⁸ *Ivi*, pp. 137-138.

²⁹ B. Petrocelli, *La colpevolezza*, cit., p. 106.

La volontà umana, essi dicono, reagisce *secondo il carattere*, e Arturo non esita ad ammettere che alla fine sia proprio sul carattere che «tutto riposa»³⁰. Il carattere certo «non è fatale»: «è una formazione complessa, in parte originaria, ma in parte acquisita», che può trasformarsi e modificarsi, anche automodificarsi, dice Rocco³¹. Perciò la determinazione psicologica della volontà non è in fondo che autodeterminazione. Ma in essa, si noti, per gli estensori del codice non c'è proprio alcun «mistero», come invece immaginava Petrocelli, nessun *fiat* che miracolosamente infranga l'ordine causale al quale, come ogni altra cosa, la volontà è soggetta³².

Nella prospettiva della «filosofia giuridica generale del Fascismo», di cui il codice «non è che una derivazione»³³, non costituiva del resto un problema il fatto che potesse essere punito chi, pur essendosi determinato ad agire per motivi coscienti secondo il suo carattere, non avrebbe potuto determinarsi ad agire diversamente, perché tale filosofia

in verità, è ben diversa da quella che fu propria degli enciclopedisti francesi a cui si ispirarono la rivoluzione del 1789, la dichiarazione dei diritti dell'uomo e del cittadino, i codici penali del 1791 e del 3 brumaio anno IV e la stessa codificazione legislativa penale del primo Napoleone, che fu il modello a cui si attenero, per più di un secolo, tutte le altre successive legislazioni che dominarono all'estero e in Italia, non escluso il codice penale del 1889³⁴.

Piacesse o meno a Petrocelli, la «generale filosofia sociale, politica e giuridica del Fascismo»³⁵ non temeva che l'autore del fatto potesse essere punito anche se, dati il suo carattere e le circostanze in cui aveva scelto di commettere il reato, non avrebbe potuto compiere una scelta differente. Né, in effetti, temeva probabilmente, nonostante le rassicurazioni di

³⁰ *Lavori preparatori del codice penale e del codice di procedura penale*, cit., p. 138.

³¹ *Ibid.*

³² Cfr. *ivi*, p. 137: «Non si può concepire [...] una volontà come un *fiat* che nasce dal nulla». Per Petrocelli, d'altra parte, è proprio in questo *fiat* che sta «tutto il mistero della volontà» (B. Petrocelli, *La colpevolezza*, cit., p. 54. E a p. 61: «A un certo momento si manifesta ciò che si denomina il *fiat* della volontà, per cui il soggetto, operata la *scelta* fra i contrastanti motivi, decide nel senso dell'uno o dell'altro di essi. Come si è già detto, non è facile stabilire [...] che cosa precisamente sia questo *fiat* della volontà, e che cosa rappresenti nel mistero del nostro essere»).

³³ *Relazione a S.M. il Re del Ministro Guardasigilli*, cit., p. 4447.

³⁴ *Ibid.*

³⁵ *Ivi*, p. 4446.

Bellavista³⁶, che in questo modo implicitamente egli fosse rimproverato per quel che era o per le circostanze in cui si era trovato, perché ammetteva *apertis verbis* che, anche in linea di principio, l'individuo potesse essere strumentalizzato per il conseguimento dei fini e per la tutela degli interessi dello Stato³⁷.

Fu del resto, come nota Donini, la stessa «concreta applicazione» del codice a dimostrare «meglio di ogni teoria dogmatica come il principio di imputazione causale ne fosse una struttura portante»³⁸. E non è difficile comprendere le ragioni per le quali, in questa prospettiva, la concezione psicologico-naturalistica della colpevolezza sembrasse la più aderente al punto di vista dell'ordinamento³⁹. Se quel che conta è solamente che l'azione o l'omissione sia stata compiuta con coscienza e volontà e che il suo autore si sia determinato a compierla «per motivi consci in conformità del [suo] carattere», la colpevolezza non può che risolversi nel nesso psichico di causalità che intercorre fra l'agente e la sua azione o, più genericamente, tra l'agente e il fatto⁴⁰.

³⁶ «Concepire la colpevolezza – scriveva nel 1942 – come *colpevolezza del carattere* non è ammissibile avanti il nostro diritto positivo» (G. Bellavista, *Il problema della colpevolezza*. Estratto dagli annali del seminario giuridico della R. Università di Palermo, Palermo, 1942, p. 66. E poco oltre, alle pp. 67-68: «Avanti al nostro diritto positivo, ancora saldamente costruito sul principio che la pena è retribuzione ed ancorato a quelli che sono stati chiamati gli eterni principi etici della penalità: “la colpa vuole l'espiazione”; “non c'è pena senza colpa” – la concezione del Liszt, del Tesar e del Kollmann non può seriamente sostenersi altro che come una “caricatura” del vero concetto di colpevolezza»). Cfr. anche M. Gallo, *Il concetto unitario di colpevolezza*, Giuffrè, Milano, 1951, pp. 101 e ss.

³⁷ *Relazione a S.M. il Re del Ministro Guardasigilli*, cit., p. 4447. Cfr. G. Neppi Modona, M. Pellissero, *La politica criminale durante il fascismo*, in L. Violante (a cura di), *Storia d'Italia. Annali 12. La criminalità*, Einaudi, Torino, 1997, p. 838.

³⁸ M. Donini, *Illecito e colpevolezza nell'imputazione del reato*, Giuffrè, Milano, 1991, p. 305.

³⁹ Cfr., ad esempio, G. Bellavista, *Il problema della colpevolezza*, cit., pp. 92, 102, 106, 109, 115, 133-134, 140.

⁴⁰ Cfr. B. Petrocelli, *La colpevolezza*, cit., pp. 31-32 e M.C. del Re, *Colpevolezza e colpevolizzazione*, Jovene, Napoli, 1976, pp. 69 e ss. Per M. Gallo, *Il concetto unitario di colpevolezza*, cit., p. 22, «il fatto che una persona ponga in essere l'elemento materiale di una fattispecie criminosa non importa, normalmente, da un punto di vista strettamente giuridico, che quell'atto venga considerato come proprio del soggetto in questione se non si riconosce che esso si trova con la volontà del suo autore in una relazione tale da permettere il riferimento alla stessa come a sua causa psichica».

3. Principio di colpevolezza e libero arbitrio nell'ordinamento repubblicano

Fortunatamente, tuttavia, dal Codice Rocco e dai suoi lavori preparatori non si dovrebbero poter desumere gli orientamenti dell'ordinamento repubblicano, al quale, d'altra parte, sono serviti quarant'anni per riconoscere la rilevanza costituzionale del «principio di colpevolezza» e adeguarvi il «sistema ordinario di cui agli artt. 42, 43, 47, 59 ecc. c.p.»⁴¹.

Fu solo nel 1988 che la Corte costituzionale avvertì, come è noto, l'esigenza di chiarire che la persona umana non può essere strumentalizzata ad alcun fine, neppure a quello della prevenzione generale, e che perciò nessuno può essere considerato penalmente responsabile di atti che non abbia scelto di compiere liberamente e che non potesse controllare⁴². Lo si desume, secondo la Corte, dal primo comma dell'art. 27 della Costituzione, che deve essere letto in «stretto collegamento» con il terzo, ed è confermato dai lavori preparatori, in particolare dai «dibattiti svoltisi, nella seduta del 18 settembre 1946, presso la I sottocommissione (della “Commissione per la Costituzione”)». Come risulta dagli atti della seduta, «tutti i Costituenti [...], almeno fino a questo momento del dibattito, sostennero che la responsabilità penale personale implicava necessariamente, oltre all'elemento materiale, un requisito subiettivo e, per alcuni Costituenti, l'esistenza in particolare, della possibilità di muovere rimprovero all'agente, potendo da lui pretendersi un comportamento diverso»⁴³.

Vi sono per la verità buone ragioni per dubitare che sia ammissibile «un'idea di colpevolezza che limiti senza fondare la potestà punitiva dello Stato»⁴⁴, ma resta il fatto che, osserva la Corte, «sia nella concezione

⁴¹ Corte cost., 24 marzo 1988, n. 364. Cfr., ad esempio, M. Donini, L. Ramponi, *Il principio di colpevolezza*, cit., pp. 291 e ss.; D. Pulitanò, *Diritto penale*, Giappichelli, Torino, 2019 (VIII ed.), p. 259; O. Di Giovine, *Delitto senza castigo? Il bisogno di pena tra motivazioni razionali ed istinti emotivi*, in «Quaderno di storia del penale e della giustizia», 2021, 3, p. 255.

⁴² Si veda, ad esempio, G. Cocco, *La colpevolezza quale principio costituzionale (e il buon vecchio libero arbitrio)*, in «Penale. Diritto e Procedura», 2 luglio 2021, p. 3. In Germania un principio assai simile era stato sancito da una notissima sentenza del *Bundesgerichtshof* del 1952 (cfr. G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, in «Materiali per una storia della cultura giuridica», XXVIII, 1998, 1, p. 218; F. Viganò, *Stato di necessità e conflitto di doveri. Contributo alla teoria delle cause di giustificazione e delle scusanti*, Giuffrè, Milano, 2002, p. 554; A. Vallini, *Antiche e nuove tensioni tra colpevolezza e diritto penale artificiale*, Giappichelli, Torino, 2003, pp. 257 e ss.).

⁴³ Corte cost., 24 marzo 1988, n. 364. Nel *Resoconto sommario della seduta di mercoledì 18 settembre 1946* si veda, in particolare, alle pp. 70-71, la replica di Aldo Moro a Ottavio Mastrojanni.

⁴⁴ Corte cost., 24 marzo 1988, n. 364. Cfr. G. Fiandaca, *Considerazioni su colpevolezza e*

che considera [la colpevolezza] “fondamento”, titolo giustificativo dell'intervento punitivo dello Stato, sia nella concezione che ne accentua particolarmente la sua funzione di limite allo stesso intervento (garanzia del singolo e del funzionamento del sistema), inalterato permane il [suo] “valore” [...], la sua insostituibilità» in un ordinamento che «pone al vertice della scala dei valori la persona umana» e vieta che per qualunque ragione e a qualunque fine possa essere strumentalizzata⁴⁵. Come la stessa Corte avrebbe chiarito nel 2007, «punire in difetto di colpevolezza, al fine di “dissuadere” i consociati dal porre in essere le condotte vietate [...] o di “neutralizzare” il reo [...], implicherebbe, infatti, una strumentalizzazione dell'essere umano per contingenti obiettivi di politica criminale (sentenza n. 364 del 1988), contrastante con il principio personalistico affermato dall'art. 2 Cost.»⁴⁶.

La Costituzione, secondo la Corte, consente di punire soltanto chi possa essere rimproverato per la commissione dell'illecito, e non si vede su quali basi possa essere rimproverata una persona per il suo comportamento se da lei non si sarebbe potuto pretendere un comportamento differente⁴⁷.

Questo non significa, precisa la Corte, che la Costituzione accolga «una delle molteplici concezioni “normative” della colpevolezza prospettate in dottrina, bensì costituisce autonomo risultato, svincolato da ogni premessa concettualistica, dell'interpretazione dei commi primo e terzo dell'art. 27 Cost., anche se, *per accidens*, il “rimprovero” venga a coincidere con una delle nozioni di colpevolezza (normativa) prospettate in dottrina»⁴⁸. Ma è difficile negare che, come hanno constatato Massimo Donini e Luca Ramponi, «al di là dei condizionamenti istituzionali che precludevano di sposare una teoria anziché un'altra», – di fatto [...] – [la Corte abbia]

prevenzione, in «Rivista italiana di diritto e procedura penale», XXX, 1987, p. 839 e, più recentemente, G. Cocco, *La colpevolezza quale principio costituzionale (e il buon, vecchio libero arbitrio)*, cit., pp. 7 e ss. Benché la Corte, pur sollevando la questione, scelga in realtà di tralasciarla, la sua importanza non può essere del tutto trascurata, perché ove si ritenesse che la colpevolezza non fondi la punibilità ma la limiti soltanto, «dovrebbe concludersi – come ha osservato Donini – che la pena è già fondata [...] dalla lesione materiale di un interesse, dalla... responsabilità oggettiva pura». Chi consideri questa conclusione inaccettabile dovrà allora convenire che «la ‘colpevolezza’ non limita la responsabilità, ma le assicura il livello minimo di consenso culturale e civile, il (solo) diritto di cittadinanza penalistica» (M. Donini, *Illecito e colpevolezza nell'imputazione del reato*, cit., p. 557). Di diverso avviso F. Viganò, *Stato di necessità e conflitto di doveri*, cit., p. 207.

⁴⁵ Corte cost., 24 marzo 1988, n. 364.

⁴⁶ Corte cost., 24 luglio 2007, n. 322.

⁴⁷ Cfr. M. Donini, L. Ramponi, *Il principio di colpevolezza*, cit., pp. 335-336.

⁴⁸ Corte cost., 24 marzo 1988, n. 364.

adottato i postulati essenziali delle concezioni normative⁴⁹, il cui carattere specifico, osservava d'altra parte Padovani nel 1973, «non consiste tanto nella prospettazione di un atteggiamento antidoveroso come punto di riferimento di tutti i coefficienti psichici della colpevolezza, intesi come oggetto del giudizio di rimproverabilità», ma «piuttosto nell'utilizzazione concreta di tale giudizio per individualizzare, ai fini dell'applicazione della pena in chiave retributiva, il rapporto tra il fatto ed il suo autore», ossia nell'«esigenza di “personalizzazione”» che le ispira⁵⁰.

Secondo tali concezioni, le cui prime elaborazioni risalgono agli inizi del secolo XX e sono legate «principalmente ai nomi del Frank, del Freudenthal, del Goldschmidt e del Mezger»⁵¹, perché qualcuno possa essere punito non basta che il fatto criminoso gli appartenga psicologicamente⁵², ma si richiede anche che la sua volontà abbia assunto al riguardo un «atteggiamento antidoveroso [...] che era possibile non assumere», e che perciò gli possa essere rimproverato⁵³.

Anche secondo le concezioni normative, come per quelle psicologiche, serve, perché vi sia colpevolezza, «un nesso psichico tra il fatto e l'autore, cioè [...] una “determinazione psicologica” assunta dal soggetto nei confronti

⁴⁹ M. Donini, L. Ramponi, *Il principio di colpevolezza*, cit., p. 335; cfr. anche pp. 300 e ss.

⁵⁰ T. Padovani, *Appunti sull'evoluzione del concetto di colpevolezza*, in «Rivista italiana di diritto e procedura penale», XVI, 1973, p. 559.

⁵¹ G. Bellavista, *Il problema della colpevolezza*, cit., pp. 44-45. Si vedano tuttavia H. Achenbach, *Riflessioni storico-domatiche sulla concezione della colpevolezza di Reinhard Frank*, trad. it. di S. Moccia, in «Rivista italiana di diritto e procedura penale», XXIV, 1981, pp. 848 e ss. e R. Bartoli, *Colpevolezza: tra personalismo e prevenzione*, Giappichelli, Torino, 2005, p. 53, secondo il quale, in particolare, «la concezione di Frank si pone, a rigore, fuori dalle concezioni normative propriamente dette».

⁵² F. Mantovani, *Diritto penale*, Cedam, Padova, 2009 (6^a ed.), p. 278.

⁵³ Ivi, p. 281. Come si legge in un'altra celebre sentenza della Corte costituzionale, «perché l'art. 27, primo comma, Cost. sia pienamente rispettato e la responsabilità penale sia autenticamente personale, è indispensabile che tutti e ciascuno degli elementi che concorrono a contrassegnare il disvalore della fattispecie siano soggettivamente collegati all'agente (siano, cioè, investiti dal dolo o dalla colpa) ed è altresì indispensabile che tutti e ciascuno dei predetti elementi siano allo stesso agente rimproverabili» (Corte cost., 13 dicembre 1988, n. 1085). La concezione normativa, ha scritto Fiandaca ormai qualche anno fa, «non esaurisce la colpevolezza nella relazione fatto-autore in forma di dolo o colpa: essa pretende di assumere questa stessa relazione ad oggetto di valutazione in termini di rimproverabilità da parte dell'ordinamento giuridico. In altri termini, così concepita la colpevolezza comporta la possibilità di rivolgere un rimprovero o biasimo all'autore per l'atteggiamento “antidoveroso” manifestato nell'aver voluto commettere un fatto criminoso che non doveva volere» (G. Fiandaca, *Prima lezione di diritto penale*, Laterza, Roma-Bari, 2017, p. 106. Cfr. R. Bartoli, *Colpevolezza*, cit., pp. 46 e ss., al quale si rinvia anche per gli ulteriori necessari riferimenti bibliografici).

del fatto»⁵⁴. Per le concezioni normative, tuttavia, la colpevolezza sussiste solamente quando si possa anche «ipotizzare se non accertare l'“autonomia” della determinazione psicologica, cioè a dire la possibilità – nella data situazione – di una diversa scelta comportamentale del soggetto»⁵⁵.

Siccome queste due componenti della colpevolezza corrispondono alle condizioni che generalmente si reputano necessarie perché un'azione possa ritenersi compiuta esercitando il libero arbitrio⁵⁶, sembrerebbe ragionevole concludere che in base alle concezioni normative della colpevolezza, nessuno potrebbe essere punito se il libero arbitrio non fosse che un'illusione⁵⁷.

⁵⁴ F. Palazzo, *Corso di diritto penale. Parte generale*, Giappichelli, Torino, 2005, p. 15. Cfr., ad esempio, E. Mezger, *Diritto penale (Strafrecht)*, trad. it. di F. Mandalari effettuata sulla seconda edizione tedesca con aggiunte dell'autore e note del traduttore in riferimento al diritto italiano, Cedam, Padova, 1935, p. 288.

⁵⁵ F. Palazzo, *Corso di diritto penale. Parte generale*, cit., pp. 15-16. Cfr., ad esempio, A. Manna, *L'imputabilità e i nuovi modelli di sanzione*, cit., p. 32.

⁵⁶ Come ha scritto Mario De Caro, «perché si possa dire correttamente che una determinata azione è compiuta in virtù del nostro libero arbitrio devono darsi due condizioni. In primo luogo, all'agente si devono presentare due o più corsi d'azione alternativi e, in secondo luogo, deve essere proprio quell'agente (e non un altro e neppure il caso) a determinare, per mezzo di una deliberazione non imposta, quale di questi corsi d'azione alternativi si attualizzerà» (M. De Caro, *Il problema filosofico della responsabilità*, in M. De Caro, A. Lavazza, G. Sartori (a cura di), *Quanto siamo responsabili? Filosofia, neuroscienze e società*, Codice, Torino, 2013, pp. 27-28. Cfr. anche M. De Caro, *Lenigma della libertà*, in Id. (a cura di), *La logica della libertà*, Meltemi, Roma, 2002, p. 9; Id., *Il libero arbitrio. Una introduzione*, Laterza, Roma-Bari, 2004, pp. 9-10; Id., *Azione*, il Mulino, Bologna, 2008, pp. 40-41; S.F. Magni, *Teorie della libertà. La discussione contemporanea*, Carocci, Roma, 2005, p. 47; R. De Monticelli, *Che cos'è una scelta? Fenomenologia e neurobiologia*, in M. De Caro, A. Lavazza, G. Sartori (a cura di), *Siamo davvero liberi? Le neuroscienze e il mistero del libero arbitrio*, Codice, Torino, 2010, p. 111; L. Sesta, *La filosofia e le chiavi della prigione. Libero arbitrio e imputabilità fra neuroscienze e diritto*, in «Diritto & Questioni pubbliche», XIX, 2019, 2, p. 246).

⁵⁷ Cfr. ad esempio W. Hassemer, *Principio di colpevolezza e struttura del reato*, in «Archivio penale», 1982, p. 54: «La disputa sul concetto di colpevolezza riguarda oggi essenzialmente i presupposti in presenza dei quali si può muovere un rimprovero di colpevolezza nei confronti di un individuo e la possibilità stessa che tali presupposti ricorrano nell'ambito del diritto penale. [...] Ciò equivale a porre la questione della libertà del volere. [...] Senza la libertà del volere nessun comportamento alternativo, senza comportamento alternativo nessun rimprovero di colpevolezza». Più in generale A. Lavazza, *Responsabilità sotto attacco e saggezza del diritto*, in M. De Caro, A. Lavazza e G. Sartori (a cura di), *Quanto siamo responsabili? Filosofia, neuroscienze e società*, cit., p. 227: «classicamente il concetto di responsabilità assunto dal diritto presuppone il libero arbitrio. Senza che un individuo agente potesse fare altrimenti, senza cioè che fosse a lui disponibile l'opzione, variamente qualificabile, di astenersi dalla condotta in questione o di intraprenderne una diversa [...] non ha senso attribuire la “colpa” dell'azione al soggetto, benché egli ne sia effettivamente la

Ma non solo: se per le concezioni normative la colpevolezza implica la «possibilità – nella data situazione – di una diversa scelta comportamentale del soggetto», nella loro prospettiva – come del resto in quella che la nostra Costituzione, secondo la Corte, avrebbe fatto propria – difficilmente essa sembrerebbe poter «trovare la propria giustificazione nella teoria del reato indipendentemente da una opzione di fondo tra determinismo e indeterminismo»⁵⁸.

“causa materiale”. [...] Tale idea è da tempo chiara ai legislatori ed è incorporata nei codici». Cfr. anche, in questo volume, F. Corso, A. Lavazza, *Diritto e neuroscienze: un avvicinamento a due velocità*, pp. 37 e ss.

⁵⁸ F. Mantovani, *Diritto penale*, cit., p. 287. Cfr. M. Bertolino, *Il breve cammino del vizio di mente. Un ritorno al paradigma organicistico?*, in A. Santosuosso (a cura di), *Le neuroscienze e il diritto*, Ibis, Pavia, 2009, p. 135: «fintanto che il modello antideterministico non verrà falsificato, anche il diritto penale dovrà continuare a rimanere fedele al principio costituzionale dell'uomo come idoneo destinatario del precetto penale, in quanto dotato della capacità di autodeterminarsi»; I. Merzagora Betsos, *Colpevoli si nasce? Criminologia, determinismo, neuroscienze*, Cortina, Milano, 2012, pp. 9-10: «il diritto penale si pone il problema della libertà del volere, e potremmo quasi dire che parte da un'opzione non deterministica (benché con significativi distinguo)». Sembrerebbero di diverso avviso Giovanni Fiandaca ed Enzo Musco – «il moderno diritto penale» non presuppone, a loro avviso, il riconoscimento della «libertà del volere» in senso filosofico» ma «si appaga piuttosto» d'una diversa libertà del volere, *relativa e condizionata*, e dunque tale da presentare «graduazioni diverse in funzione del livello di intensità dei condizionamenti, anche di natura inconscia, che il soggetto subisce prima di agire» (G. Fiandaca, E. Musco, *Diritto penale. Parte generale*, Zanichelli, Bologna, 2019, 6ª ed., pp. 346-347) – e Domenico Pulitanò, il quale constata che «*le idee di responsabilità, colpevolezza e pena, nel nostro orizzonte di civiltà, sono legate all'idea dell'uomo come soggetto morale, capace di autodeterminarsi liberamente, nel bene e nel male*», ma ritiene che «ciò non signific[hi] postulare una volontà non soggetta a condizionamenti, né una presa di posizione sul problema filosofico del libero arbitrio. La capacità di autodeterminarsi, che il diritto presuppone, è la capacità che in via di principio postuliamo nell'uomo adulto, a partire da una certa età, e che può venir meno in condizioni particolari, per es. per infermità psichica o fisica. [...] Quando parliamo di capacità di autodeterminazione o di volontà libera, non pensiamo a scelte effettuate senza ragioni o cause esterne, ma alla capacità di controllare gli impulsi, di resistere a condizionamenti, di scegliere in base a ragioni» (D. Pulitanò, *Diritto penale*, cit., pp. 329-330). Assai simili le conclusioni di O. Di Giovine, *Prove di dialogo tra neuroscienze e diritto*, in «Giornale italiano di psicologia», 4, 2016, p. 720 e C. Grandi, *Neuroscienze e responsabilità penale. Nuove soluzioni per antichi problemi?*, Giappichelli, Torino, 2016, pp. 107, 110 e 118. Cfr. tuttavia P.F. Strawson, *Freedom and Resentment*, in «Proceedings of the British Academy», 48, 1962, trad. it., *Libertà e risentimento*, in M. De Caro (a cura di), *La logica della libertà*, cit., p. 106: «Quale sia il senso di “determinato” necessario ad asserire la tesi del determinismo, è poco probabile che sia tale da permettere compromessi o risposte-limite alla domanda “Questo esempio di comportamento è determinato oppure no?”». Come ha osservato più recentemente S.J. Morse, *New Neuroscience, Old Problems*, cit., p. 174: «determinism is not a degree or continuum concept. To say, as many do, that there is a continuum between

4. Una prima difficoltà, d'ordine logico

La prima e la più cospicua difficoltà teorica che presentano le concezioni normative sembra prescindere, in realtà, da qualsivoglia ipotesi sul mondo e ha piuttosto natura logica e concettuale.

Se si concepisce deterministicamente il nesso psichico che costituisce la prima componente della colpevolezza, non si vede come tale nesso potrebbe sussistere laddove fossero riscontrati i «presupposti» e le «condizioni in presenza delle quali si può quantomeno ipotizzare se non accertare l'«autonomia» della determinazione psicologica, cioè a dire la possibilità – nella data situazione – di una diversa scelta comportamentale del soggetto»⁵⁹. Se non lo si concepisce deterministicamente, si deve ammettere che all'agente possa essere rimproverato un atteggiamento della volontà che gli è capitato di assumere in modo del tutto casuale.

L'argomento è ben noto al dibattito sul libero arbitrio⁶⁰, il cui concetto secondo Hans Vaihinger «non contraddice soltanto la realtà osservata, [...] ma contraddice anche se stesso», poiché «una condotta assolutamente

determinism and free will is simply a conceptual error».

⁵⁹ F. Palazzo, *Corso di diritto penale*, cit., pp. 15-16.

⁶⁰ La letteratura al riguardo è realmente sterminata. Cfr., ad esempio, G. Ryle, *The Concept of Mind*, Barnes and Noble, New York, 1949, trad. it. di G. Pellegrino, *Il concetto di mente*, Laterza, Roma-Bari, 2007, pp. 77 e ss.; P.H. Nowell-Smith, *Ethics*, Penguin Books, London, 1954, p. 28; A.J. Ayer, *Freedom and Necessity*, in Id., *Philosophical Essays*, Macmillan, London-Basingstoke, 1954, trad. it., *Libertà e necessità*, in M. De Caro (a cura di), *La logica della libertà*, cit., p. 45; R.M. Chisholm, *Human Freedom and the Self: The Lindley Lecture*, Department of Philosophy, University of Kansas, 1964, trad. it., *La libertà umana e il sé*, in M. De Caro (a cura di), *La logica della libertà*, cit., pp. 60-61; A. Kenny, *Freewill and Responsibility*, Routledge and Kegan Paul, London, 1978, p. 26; R. Nozick, *Philosophical Explanations*, Harvard University Press, Cambridge (Mass.), 1981, trad. it. di G. Rigamonti, *Spiegazioni filosofiche*, Il Saggiatore, Milano, 1987, pp. 332 e 339; T. Nagel, *The View from Nowhere*, Oxford University Press, New York, 1986, trad. it. di A. Besussi, *Uno sguardo da nessun luogo*, Il Saggiatore, Milano, 1988, p. 141; S. Wolf, *Freedom within Reason*, Oxford University Press, New York, 1990, p. 13; R. Kane, *Responsibility, Luck and Chance: Reflections on Free Will and Indeterminism*, in «Journal of Philosophy», 96, 5, 1999, p. 219 e Id., *Libertarianism*, in J.M. Fischer, R. Kane, M. Pereboom e M. Vargas, *Four Views on Free Will*, Blackwell, Malden (Mass.), 2007, p. 9; P. Van Inwagen, *The Mystery of Metaphysical Freedom*, in P. Van Inwagen, D.W. Zimmerman (eds.), *Metaphysics: The Big Questions*, Blackwell, Malden (Mass.), 1998, trad. it., *Il mistero della libertà metafisica*, in M. De Caro (a cura di), *La logica della libertà*, cit., p. 167; J.D. Greene, J. Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, cit., p. 211; A.R. Mele, *Free Will and Luck*, Oxford University Press, New York, 2006, p. 14; M. Balaguer, *Free Will*, The MIT Press, Cambridge (Mass.)-London, 2014, p. 80.

libera, assolutamente non determinata, e che pertanto non segua da nulla è, dal punto di vista morale, altrettanto priva di valore quanto una condotta assolutamente necessaria»⁶¹.

Perché si possa ragionevolmente sostenere che, *nella situazione data*, l'agente avrebbe potuto compiere una scelta diversa da quella che ha compiuto, è necessario assumere che quest'ultima non sia stata determinata da alcuna causa, dunque neppure da colui che l'ha compiuta⁶². Come ha osservato Dennett: «se la decisione è indeterminata [...], non è determinata da voi, qualunque cosa voi siate, perché non è determinata da niente»⁶³. Laddove si riuscisse perciò a provare l'esistenza di un nesso psichico adeguato fra l'autore e il fatto, ossia che la scelta di commettere il reato sia stata «l'effetto di una catena causale in cui i desideri, le credenze e le intenzioni dell'agente [hanno giocato] un ruolo determinante»⁶⁴, si dovrebbe escludere che, trovandosi nel medesimo stato e nelle stesse circostanze, l'agente avrebbe potuto scegliere di non commettere il reato.

Per converso, laddove si ritenessero sussistere i «presupposti» e le «condizioni in presenza delle quali si può quantomeno ipotizzare se non accertare [...] la possibilità – nella data situazione – di una diversa scelta comportamentale del soggetto»⁶⁵, si dovrebbe escludere che sia intercorso fra l'agente e i 'suoi' atti un nesso tale da consentire di ritenere che effettivamente fossero sotto il suo controllo. E, come afferma la nostra Corte costituzionale, «il principio di colpevolezza è [...] indispensabile [...] per garantire al privato [...] che sarà chiamato a rispondere penalmente solo per azioni da lui controllabili»⁶⁶.

Stando così le cose, la conclusione che se ne dovrebbe trarre è che mai nessuno dovrebbe potersi ritenere colpevole di alcun reato⁶⁷.

⁶¹ H. Vaihinger, *Die Philosophie des Als Ob: System der theoretischen, praktischen und religiösen Fiktionen der Menschheit auf Grund eines idealistischen Positivismus* (1911), Meiner, Leipzig, 1922, trad. it. di F. Voltaggio, *La filosofia del "come se". Sistema delle finzioni scientifiche, etico-pratiche e religiose del genere umano*, Ubaldi, Roma, 1967, p. 53.

⁶² M. De Caro, *Lenigma della libertà*, cit., pp. 12 e 25; Id., *Il libero arbitrio*, cit., pp. 19 e 127; Id., *Azione*, cit., p. 56.

⁶³ D.C. Dennett, *L'evoluzione della libertà*, cit., p. 165.

⁶⁴ M. De Caro, *Il libero arbitrio*, cit., p. 10.

⁶⁵ F. Palazzo, *Corso di diritto penale*, cit., pp. 15-16.

⁶⁶ Corte cost., 24 marzo 1988, n. 364.

⁶⁷ Riprendo qui alcune osservazioni fatte in L. Milazzo, *Liberi tutti? Alcune considerazioni su libero arbitrio e colpevolezza*, Giappichelli, Torino, 2018, pp. 80-81 e 89-90, rinviando ai capitoli V-VIII per una più diffusa trattazione del problema.

5. Dalla libertà del volere alla «motivabilità mediante norme»

Le ragioni che hanno indotto dottrina e giurisprudenza (compresa quella costituzionale) a svuotare pressoché del tutto, subito dopo averlo proclamato, il principio di colpevolezza della sua funzione garantista per farne il fondamento ideologico fittizio della potestà punitiva dello Stato, tuttavia, non sono queste, ed è probabile che debbano essere cercate innanzitutto nella consapevolezza del fatto che, se davvero si potesse punire qualcuno solo dopo aver *provato* che avrebbe potuto assumere un diverso atteggiamento della volontà, nessuno in alcun caso potrebbe essere punito.

È ormai pacifico in letteratura che «una colpevolezza [...] intesa come possibilità individuale di agire altrimenti all'epoca del fatto non è suscettibile di accertamento». Pertanto, «se [...] per affermare la colpevolezza si presuppone una prova di tipo empirico, che per principio non può essere acquisita, ciò dovrebbe condurre ad una pronuncia assolutoria, secondo il principio “*in dubio pro reo*”»⁶⁸.

Alcuni immaginano che il problema possa essere risolto riducendo il principio di colpevolezza, il quale vieta di punire chi non avrebbe potuto assumere un diverso atteggiamento della volontà *nella situazione data*, ad «un principio regolativo di tipo giuridico»⁶⁹ che prescrive di trattare l'autore *come se* fosse stato libero di scegliere di agire diversamente «quando si possa accertare che [...], all'epoca della commissione del fatto, fosse fondamentalmente motivabile mediante norme»⁷⁰. Benché infatti «la

⁶⁸ C. Roxin, *Sul problema del diritto penale della colpevolezza*, in «Rivista italiana di diritto e procedura penale», XXVII, 1984, p. 18. Si veda anche, ivi, p. 24 e W. Hassemer, *Principio di colpevolezza e struttura del reato*, cit., p. 59: «Dire che qualcuno ha mancato la possibilità di comportarsi diversamente da come si è comportato, significa presupporre che egli avesse di fatto tale possibilità alternativa. La relativa prova non si può insomma fornire senza l'indicazione delle alternative effettivamente esistenti che si offrivano all'agente nella situazione concreta». Cfr. B. Schünemann, *L'evoluzione della teoria della colpevolezza nella repubblica federale tedesca*, in «Rivista italiana di diritto e procedura penale», XXXIII, 1990, p. 4; F. Viganò, *Stato di necessità e conflitto di doveri*, cit., pp. 557 e ss.; E. Venafro, *Scusanti*, Giappichelli, Torino, 2002, pp. 84-85; R. Bartoli, *Colpevolezza*, cit., p. 65.

⁶⁹ C. Roxin, *Sul problema del diritto penale della colpevolezza*, cit., p. 24.

⁷⁰ Ivi, p. 27. Si vedano al riguardo H. Achenbach, *Individuelle Zurechnung, Verantwortlichkeit, Schuld*, in B. Schünemann (Ed.), *Grundfragen des Modernen Strafrechtssystems*, de Gruyter, Berlin-New York, 1984, p. 149 e B. Schünemann, *Die Funktion des Schuldprinzips in Präventionsstrafrecht*, in B. Schünemann (Ed.), *Grundfragen des Modernen Strafrechtssystems*, cit., pp. 160-161. Secondo Schünemann in questo modo il concetto di colpevolezza sarebbe ridotto a una «finzione» (*Fiktion*), la quale peraltro non opererebbe affatto, in realtà, come vorrebbe Roxin, sempre soltanto a favore del reo e mai suo danno. Cfr. T. Padovani,

motivabilità mediante norme [...] [sia] indipendente dalla sussistenza della libertà del volere nell'uomo e così pure dal fatto se l'agente, nel suo concreto spazio decisionale, avrebbe effettivamente potuto determinarsi altrimenti, [...] è l'unica cosa che [...] possa essere empiricamente accertata»⁷¹.

Si tratta, in apparenza, di una soluzione convincente: poiché il diritto intende determinare mediante norme le condotte dei consociati ponendosi come «*fattore motivante* il comportamento umano», quando «il precetto penale non è in grado di poter funzionare come fattore motivante, [...] allora *non avrebbe senso* irrogare una pena»⁷².

A ben vedere, tuttavia, le ragioni per le quali sarebbe insensato punire chi non avrebbe potuto essere motivato dalla norma penale a non commettere il reato sono tutt'altro che evidenti. Non è scontato affatto, innanzitutto – non lo è per tutti, almeno – che farlo sia sempre necessariamente inutile o dannoso⁷³. Come ha osservato Fiandaca, le

Teoria della colpevolezza e scopi della pena. Osservazioni e rilievi sui rapporti tra colpevolezza e prevenzione con riferimento al pensiero di Claus Roxin, in «Rivista italiana di diritto e procedura penale», XXX, 1987, p. 825; F. Palazzo, *Corso di diritto penale*, cit., pp. 405-407; R. Bartoli, *Colpevolezza*, cit., pp. 70 e ss.; M. Donini, L. Ramponi, *Il principio di colpevolezza*, cit., p. 344. Cfr. anche C. Castelfranchi, *Neuro-Norme: per un approccio non riduzionista. Cosa cercare e non cercare nel cervello*, in «Rivista di filosofia del diritto», numero speciale *Diritto e neuroscienze*, 3, 2014, p. 24 e S. Zullo, *Causalità naturale e responsabilità. "Colpa o merito" delle neuroscienze?*, in R. Brighi, S. Zullo (a cura di), *Filosofia del diritto e nuove tecnologie*, cit., p. 163.

⁷¹ C. Roxin, *Sul problema del diritto penale della colpevolezza*, cit., p. 26. Cfr. F. Viganò, *Stato di necessità e conflitto di doveri*, cit., pp. 557 e ss.

⁷² F. Palazzo, *Corso di diritto penale*, cit., pp. 406-407. Cfr. anche W. Hassemer, *Principio di colpevolezza e struttura del reato*, cit., p. 62; T. Padovani, *Teoria della colpevolezza e scopi della pena*, cit., p. 812 e, soprattutto, pp. 819 e ss.; G. Fiandaca, E. Musco, *Diritto penale*, cit., p. 348; G. De Francesco, *Diritto penale. I fondamenti*, Giappichelli, Torino, 2011, p. 366 e p. 373; E. Venafro, *Scusanti*, cit., pp. 111 e ss.; C. Grandi, *Neuroscienze e responsabilità penale*, cit., p. 124.

⁷³ Per Hart, ad esempio, il ben noto argomento benthamiano secondo il quale ci si astiene in genere dal punire i minori, i malati di mente e gli ubriachi perché, per una ragione o per l'altra, non si trovano nella «disposizione d'animo in cui la prospettiva di mali così distanti come quelli enunciati dalle leggi possa avere l'effetto di influenzare la [loro] condotta» in realtà non è che «uno spettacolare *non sequitur*» (J. Bentham, *An Introduction to the Principles of Morals and Legislation*, Payne, London, 1789, trad. it. e note di S. Di Pietro, *Introduzione ai principi della morale e della legislazione*, a cura di E. Lecaldano, Utet, Torino, 1998, XIII, V, 9, p. 278 e H.L.A. Hart, *Prolegomenon to the Principles of Punishment*, in «Proceedings of the Aristotelian Society», 60, 1959 e in Id., *Punishment and Responsibility. Essays in the Philosophy of Law* (1968), Oxford, Oxford University Press, 2008 (2^a ed.), trad. it. di M. Jori, *Introduzione ai principi della pena*, in Id., *Responsabilità e pena*, Edizioni di Comunità, Milano 1981, p. 45). Bentham «si propone di provare che punire il pazzo o

conoscenze delle quali disponiamo non consentono di «sostenere che la punizione di fatti incolpevoli mai [sia] in grado di sortire un'efficacia deterrente»⁷⁴ e ancora «le scienze psicologiche sono ben lungi dal fornire prova empirica della validità della tesi che vuole gli incapaci insuscettivi di motivazione mediante minacce di pena»⁷⁵.

Ma anche ammesso che punire chi non avrebbe potuto essere motivato dalla norma non possa soddisfare in alcun caso le esigenze della prevenzione speciale e generale, quantomeno ove quest'ultima sia intesa in senso positivo, resta comunque il fatto che, risolvendo la possibilità di determinarsi altrimenti in quella di essere motivati dalle norme dell'ordinamento si orienta teleologicamente il concetto della colpevolezza fino a privarla quasi del tutto della sua «funzione di 'argine' [...] nei confronti di un'espansione, altrimenti incontrollata (e incontrollabile) della potestà punitiva statutale»⁷⁶,

l'infante o quanti violino il diritto non intenzionalmente, o spinti da minacce o perfino in stato di 'necessità', debba essere inefficace; ma tutto quello che prova (al meglio) è la proposizione assai differente che la minaccia della pena sarà inefficace per quanto riguarda la classe di persone che sono afflitte da queste condizioni» (H.L.A. Hart, *Murder and the Principles of Punishment: England and the United States*, in «Northwestern University Law Review», 52, 4, 1957, nonché in Id., *Punishment and Responsibility*, cit., trad. it. di M. Jori, *Omicidio e principi della pena: Inghilterra e Stati Uniti*, in Id., *Responsabilità e pena*, cit., p. 103). Cfr. anche G. Fiandaca, *Considerazioni su colpevolezza e prevenzione*, cit., p. 843 e F. Viganò, *Stato di necessità e conflitto di doveri*, cit., p. 560.

⁷⁴ G. Fiandaca, *Considerazioni su colpevolezza e prevenzione*, cit., p. 842.

⁷⁵ Ivi, p. 843. Cfr. anche G. Fiandaca, E. Musco, *Diritto penale. Parte generale*, cit., pp. 340. Si vedano, tuttavia, H. Achenbach, *Individuelle Zurechnung, Verantwortlichkeit, Schuld*, cit., pp. 143 e ss. e, più recentemente, R. Bartoli, *Colpevolezza*, cit., pp. 29 e ss.

⁷⁶ T. Padovani, *Teoria della colpevolezza e scopi della pena*, cit., p. 799; cfr. anche pp. 813 e 822. Come ha ricordato di recente Dante Valitutti, la «tendenza già ben presente nello sviluppo della penalistica [...] a subordinare la colpevolezza agli scopi della prevenzione» si «radicalizza» nel «sistema dell'imputazione» di Günther Jakobs fino a condurre ad «un 'asservimento' totale [...] della colpevolezza alla prevenzione» (D. Valitutti, *Normativismo e funzionalismo penale. Saggio sulla teoria giuridica di Günther Jakobs*, Giappichelli, Torino, 2020, pp. 147-148 e, ivi, pp. 191 e ss. Cfr., ad esempio, G. Jakobs, *La funzione del dolo, della colpa e della colpevolezza nel diritto penale*, in *Studi sulla colpevolezza*, a cura di L. Mazza, Giappichelli, Torino, 1990, pp. 15-43). Secondo Fiandaca, «quanto più la categoria della colpevolezza è co-determinata da preoccupazioni finalistiche, tanto meno essa è in grado di controbilanciare le esigenze preventive» (G. Fiandaca, *Considerazioni su colpevolezza e prevenzione*, cit., p. 859). Come ha scritto Hassemer, «il diritto penale non deve adattarsi alle esigenze della politica criminale, non deve diventare un sottosistema (funzionale) ma deve al contrario opporsi ad esso, deve frenarla e se necessario metterle i bastoni fra le ruote» (W. Hassemer, *Principio di colpevolezza e struttura del reato*, cit., p. 6). E G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, cit., p. 224: «l'imputazione di colpevolezza segue leggi sue proprie, le quali non possono essere adeguatamente riformulate nelle categorie funzionali della politica criminale» (cfr. anche,

senza riuscire peraltro, in questo modo, ad affrancarla dai suoi presunti «condizionamenti metafisici»⁷⁷.

Illustrando uno dei molti tentativi compiuti per «determinare un fondamento psicologico della responsabilità penale diverso dal libero arbitrio», nel 1937 Scarano ricordava in breve i tratti della cosiddetta «teoria della intimidabilità»⁷⁸. Secondo questa teoria, che a suo avviso «aveva ricevuto la sua maggiore precisazione dall'Impallomeni» e «aveva avuto il suo precursore nello Stuart Mill»⁷⁹, ma che in realtà si può molto probabilmente far risalire almeno a Bentham⁸⁰, «perché l'agente sia imputabile, si presuppone che vi sia possibilità, al tempo del reato, che la pena, inserendosi nella vita della coscienza, agisca quale contropinta all'impulso criminoso»⁸¹. Poiché la norma penale ha la funzione specifica di prevenire i reati tramite l'intimidazione, «la intimidabilità della maggior parte degli uomini sani e maturi di mente – scrive Impallomeni – è [...] la proprietà psicologica che dà ragione della imputabilità»⁸².

Scarano esclude, tuttavia, che questa teoria sia effettivamente in grado di dare conto della responsabilità penale indipendentemente da qualsivoglia opzione riguardo alla verità del determinismo.

Come aveva osservato Enrico Ferri, «se un uomo commette un delitto gli è precisamente perché egli non fu intimidito e – nelle precise condizioni di fatto in cui agì – non era intimidabile dalla pena! Sicché la conseguenza logica di questa teoria sarebbe che responsabili sarebbero soltanto gli uomini che non commettono delitti»⁸³. Per chi creda, come Ferri, che sia

ivi, pp. 25 e ss. e F. Viganò, *Stato di necessità e conflitto di doveri*, cit., p. 571). Secondo Pulitanò, per altro verso, «la contrapposizione dei punti di vista funzionale e garantista non può essere rappresentata come opposizione fra un'esigenza di limitazione ed un'esigenza di espansione dell'intervento penale. Anche la prospettiva funzionale pone un vincolo interno di razionalità strumentale, e problemi di verifica della razionalità degli istituti e delle prassi di attribuzione della responsabilità» (D. Pulitanò, *Diritto penale*, cit., p. 255).

⁷⁷ F. Palazzo, *Corso di diritto penale*, cit., p. 20; M. Donini, L. Ramponi, *Il principio di colpevolezza*, cit., p. 344. Più ampiamente in L. Milazzo, *Liberi tutti?*, cit., pp. 11 e ss.

⁷⁸ L. Scarano, *Libera volontà e libero arbitrio nel diritto penale*, cit., p. 48.

⁷⁹ *Ibid.* Cfr. E. Venafro, *Scusanti*, cit., pp. 45 e ss.

⁸⁰ Cfr., *supra*, nota 72 e G. Impallomeni, *Istituzioni di diritto penale*. Opera postuma curata da V. Lanza, Unione Tipografico-Editrice Torinese, Torino, 1908, pp. 234-235.

⁸¹ L. Scarano, *Libera volontà e libero arbitrio nel diritto penale*, cit., p. 48.

⁸² G. Impallomeni, *Istituzioni di diritto penale*, cit., p. 240. Cfr. G. Fiandaca, *Considerazioni su colpevolezza e prevenzione*, cit., p. 841. Ma si veda anche G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, cit., pp. 221-222, il quale ricorda che anche «Franz von Listz aveva definito l'imputabilità [...] come "la ricettività della motivazione della pena"».

⁸³ E. Ferri, *Sociologia criminale*, Bocca, Milano-Roma-Firenze, 1900 (4^a ed.), p. 641. Cfr. L.

vero il determinismo, nessuno avrebbe mai potuto fare altro che quel che ha fatto⁸⁴. Perciò ogni qual volta qualcuno commetta un reato, per il solo fatto che lo ha commesso si dovrà escludere che, nelle stesse circostanze, avrebbe potuto (essere motivato dalla norma a) non commetterlo.

6. *Dalla possibilità di assumere l'atteggiamento doveroso della volontà alla constatazione del fatto che gli altri in genere lo fanno*

Appurato «che – nel caso concreto – rimangono indimostrati sia la possibilità del soggetto di agire altrimenti sia le ragioni per cui è prevalso il movente criminoso»⁸⁵, per alcuni sarebbe cionondimeno possibile «fondare giuridicamente un giudizio di colpevolezza attraverso la costruzione normativa di modelli di normalità psichica in relazione ai quali [sia] affermabile la possibilità di una diversa determinazione di volontà». Il «libero arbitrio», in altri termini, dovrebbe essere *presunto* «in presenza di condizioni psichiche “normali”», ossia quando «le caratteristiche della personalità del soggetto (come individuate dal diritto mediante modelli fondati sull'età, salute, ecc.) fossero tali per cui la maggioranza delle persone corrispondenti a quei modelli si astiene di regola dal commettere quel reato»⁸⁶.

Anche ammesso, dunque, che, *tutto considerato*, nessuno avrebbe potuto agire diversamente da come ha agito o che comunque non si potrebbe mai dimostrare che lo avrebbe potuto fare, si dovrebbe considerare colpevole l'autore del fatto quando sia «un uomo di predisposizione media, in condizioni medie» e non sia stato «spinto causalmente» a commettere il reato «da certe circostanze diverse da quelle medie presupposte dall'ordinamento giuridico»⁸⁷.

Fin dalle loro prime elaborazioni, del resto, le stesse concezioni normative della colpevolezza, nonostante la loro vocazione personalistica

Scarano, *Libera volontà e libero arbitrio nel diritto penale*, cit., p. 50.

⁸⁴ P. Van Inwagen, *Il mistero della libertà metafisica*, cit., p. 165. Cfr. anche A.R. Mele, *Free Will and Luck*, cit., p. 19; B. Berofsky, *Ifs, Cans, and Free Will: The Issues*, in R. Kane (ed.), *Oxford Handbook of Free Will*, Oxford University Press, Oxford, 2002, p. 198; D.C. Dennett, *L'evoluzione della libertà*, cit., pp. 100 e ss.

⁸⁵ F. Palazzo, *Corso di diritto penale*, cit., p. 16.

⁸⁶ *Ibid.* Cfr. anche E. Venafro, *Scusanti*, cit., pp. 85-86 e G. Cocco, *La colpevolezza quale principio costituzionale (e il buon, vecchio libero arbitrio)*, cit., pp. 10-11.

⁸⁷ H. Kelsen, *La dottrina pura del diritto*, cit., p. 117. Cfr. F. Viganò, *Stato di necessità e conflitto di doveri*, cit., pp. 296-297.

e individualizzante, dovettero accontentarsi, come ricorda Fornasari, di commisurare il «dovere dell'agente» alle possibilità dell'«uomo medio» (James Goldsmith)⁸⁸ o, tutt'al più, al «poter fare generale» del «cittadino medio» (Eberhard Schmidt)⁸⁹.

E come è noto, la stessa Corte costituzionale, dopo aver desunto «dal collegamento tra il primo e terzo comma dell'art. 27 Cost. [...] la necessaria «rimproverabilità» della *personale* violazione normativa»⁹⁰ e, dunque, «l'illegittimità costituzionale di fatti che non risultino essere espressione di consapevole, rimproverabile contrasto con i (od indifferenza ai) valori della convivenza espressi dalle norme penali», chiarisce che

l'inevitabilità dell'errore sul divieto (e, conseguentemente, l'esclusione della colpevolezza) non va misurata alla stregua di criteri c.d. soggettivi puri (ossia di parametri che valutino i dati influenti sulla conoscenza del precetto esclusivamente alla luce delle specifiche caratteristiche *personali* dell'agente) bensì secondo criteri oggettivi: ed anzitutto in base a criteri (c.d. oggettivi puri) secondo i quali l'errore sul precetto è inevitabile nei casi d'impossibile conoscenza della legge penale da parte d'ogni consociato⁹¹.

Soluzioni come queste, per quanto dettate prevalentemente da esigenze di natura probatoria⁹², consentono, esse sì, effettivamente di «sottrarsi

⁸⁸ G. Fornasari, *Il principio di inesigibilità nel diritto penale*, Cedam, Padova, 1990, pp. 67-68.

⁸⁹ Ivi, pp. 75 e 78. Ma Fornasari osserva anche che Berthold Freudenthal restò invece sempre dell'avviso che «l'indagine sul venir meno del potere concreto che dev'essere alla base del rimprovero di colpevolezza non possa essere generalizzata, ma vada effettuata tramite un accurato controllo della situazione di fatto del caso concreto» (G. Fornasari, *Il principio di inesigibilità nel diritto penale*, cit., p. 82. Cfr. anche F. Viganò, *Stato di necessità e conflitto di doveri*, cit., pp. 83 e ss.).

⁹⁰ Corte cost., 24 marzo 1988, n. 364. Corsivo aggiunto.

⁹¹ *Ibid.* Corsivo aggiunto.

⁹² Cfr., ad esempio, W. Hassemer, *Principio di colpevolezza e struttura del reato*, cit., pp. 59-60: «La sostituzione del criterio dell'uomo concreto con quello dell'uomo medio è la conseguenza obbligata della presa di coscienza che con i mezzi del processo penale non è dato stabilire alternative di azione di fatto esistenti. [...] Il «potere individuale» di agire in maniera diversa da come si è agito è una specie di cambiale che nel processo penale nessuno potrà mai onorare». E Stratenwerth: «Se effettivamente si vuol rimproverare al reo che avrebbe potuto comportarsi diversamente da come aveva agito, è necessario allora individualizzare rigorosamente i presupposti della colpevolezza. Non basta [...] speculare in generale sul libero arbitrio. Occorre provare piuttosto nel singolo caso che l'agente concreto, nel momento irripetibile, si era deciso in piena libertà contro il diritto. Sennonché una prova siffatta non è possibile» (G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, cit., pp. 219-220).

all'inesorabilità della critica deterministica», ma ad un prezzo veramente molto alto⁹³.

Una volta, infatti, che si sia rinunciato alla «pretesa di fondare il rimprovero di colpevolezza su un “potere individuale di agire altrimenti”»⁹⁴, all'autore del fatto che disconosca la legittimità della pena che gli è inflitta, scrive Hassemer,

non [...] si potrà in tutta coscienza rispondere altro: ciò che avresti dovuto fare, noi lo sappiamo; se tu però, nella tua situazione specifica, lo avresti potuto davvero fare, noi non lo sappiamo; né sappiamo, quindi, se tu abbia colpa o meno per il tuo distacco dalla media. Ed in fondo, per il tipo di rimprovero che ti muoviamo, non ci interessa nemmeno saperlo: ci basta il tuo distacco obiettivo dall'uomo medio, ed un tale distacco l'abbiamo già esattamente misurato⁹⁵.

Una volta che si sia rinunciato «all'individualizzazione del giudizio di colpevolezza», rileva Stratenwerth, «quest'ultimo non può più essere configurato affermando che *questo agente* avrebbe potuto agire diversamente, ma solo che “*un altro*” nella sua situazione “avrebbe potuto resistere alla tentazione del reato”»⁹⁶. E tuttavia, «chiedersi se un soggetto terzo, oppure un uomo medio, era in grado di decidere diversamente» in realtà «non ha alcun senso»⁹⁷, poiché non ha alcun senso infliggere deliberatamente dolore e sofferenza a qualcuno perché (se si preferisce: solo se) qualcun altro nella sua situazione avrebbe potuto agire diversamente⁹⁸ o, più precisamente, *perché altri, che condividono con lui alcune caratteristiche alle*

⁹³ W. Hassemer, *Principio di colpevolezza e struttura del reato*, cit., p. 57.

⁹⁴ Ivi, p. 58.

⁹⁵ Ivi, p. 67.

⁹⁶ G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, cit., p. 226.

⁹⁷ R. Bartoli, *Colpevolezza*, cit., p. 69. Cfr. anche E. Venafro, *Scusanti*, cit., pp. 86 e 151.

⁹⁸ «L'obiezione di fondo che può essere indirizzata ad un simile concetto sociale di colpevolezza, una volta spogliato dei suoi contenuti metafisici e riconsiderato pragmaticamente, è quella di non essere in grado di legittimare la pena. Anche qualora il concetto di colpevolezza non dovesse essere il fondamento dell'intervento punitivo, ma dovesse solo servire a limitarlo, non si può sfuggire al fatto che tale intervento è connesso ad un rimprovero personale che discrimina l'agente in una maniera che spesso è più temuta della pena vera e propria. [...] Senonché tale rimprovero non può essere giustificato con l'argomento per cui, in base alla nostra esperienza, un altro soggetto al posto dell'agente avrebbe potuto agire in maniera conforme al diritto. Esso si indirizza proprio contro l'agente stesso il quale da parte sua può sempre dire che egli appunto non è quell'altro» (G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, cit., pp. 227-228).

quali l'ordinamento conferisce rilevanza, in situazioni simili in genere non commettono il reato⁹⁹.

7. 'Come se'

Perciò chi teme «che il riferimento all'uomo comune medio sottragga al giudizio di colpevolezza ogni fondamento reale, con la conseguenza di trasformare la colpevolezza stessa in una categoria vuota di contenuto perché completamente asservita ad esigenze di prevenzione»¹⁰⁰, insiste sulla necessità «di valutare la capacità "individuale" di agire diversamente»¹⁰¹ e ribadisce che «il giudice [deve] accertare il potere individuale di agire altrimenti del soggetto concretamente sottoposto a giudizio, e cioè di una persona in carne ed ossa dotata di specifiche caratteristiche e attitudini»¹⁰².

Si illuderebbe, d'altra parte, chi ritenesse di poter desumere da alcuni suoi orientamenti ormai consolidati che la Corte di Cassazione condivida fino in fondo questo genere di preoccupazioni. Certo, si noterà, sono passati ormai più di tre lustri da quando la Corte ha ammesso che anche i disturbi atipici della personalità possano «escludere o scemare grandemente la capacità di intendere e di volere», chiarendo per altro che l'accertamento del nesso eziologico fra l'infermità e il reato debba avvenire «attraverso un approccio non astratto ed ipotetico, ma reale ed individualizzato, in specifico riferimento, quindi, alla stessa sfera di possibile, o meno, autodeterminazione della persona cui quello specifico fatto di reato medesimo si addebita e si rimprovera»¹⁰³.

È forse superfluo rammentare, tuttavia, che un simile sforzo probatorio non è richiesto dalla Corte all'accusa ma alla difesa, perché, come

⁹⁹ Per una più ampia trattazione del problema rinvio a L. Milazzo, *Liberi tutti?*, cit., pp. 113 e ss.

¹⁰⁰ G. Fiandaca, E. Musco, *Diritto penale*, cit., p. 342.

¹⁰¹ *Ibid.*

¹⁰² *Ibid.* Riguardo all'esigenza di individualizzazione del giudizio di colpevolezza del reato colposo si veda in particolare D. Castronuovo, *La colpa penale*, Giuffrè, Milano, 2009, pp. 597-598.

¹⁰³ Cass., Sez. Un., 25 gennaio 2005, n. 9163. Cfr. M. Bertolino, *Il breve cammino del vizio di mente*, cit., p. 128 e D. Piva, *Le componenti impulsive della condotta. Tra imputabilità, (pre)colpevolezza e pena*, Jovene, Napoli, 2020, pp. 129-130 nonché, in questo volume, *Il ruolo delle emozioni nel giudizio penale tra "oggetto della valutazione" e "valutazione dell'oggetto"*, pp. 97-98..

subito la Corte provvede a ribadire «la imputabilità è normalmente considerata presente quando l'autore abbia raggiunto la maturità psicofisica normativamente indicata [...] salvo che versi in una situazione di infermità, tanto costituendo (ancora per autorevole voce della dottrina) “il compromesso, o il punto d'incontro, tra le esigenze proprie del principio di colpevolezza e quello della prevenzione generale”»¹⁰⁴.

Ora, a prescindere dal fatto che, come la Corte costituzionale ha avuto modo di chiarire, «il principio di colpevolezza non può essere “sacrificato” dal legislatore ordinario in nome di una più efficace tutela penale di altri valori, ancorché essi pure di rango costituzionale»¹⁰⁵, non si può fare a meno di notare che la stessa Corte nel 1988 aveva operato un analogo paradossale ribaltamento *in malam partem* del principio personalistico, prescrivendo che delle peculiari competenze dell'agente si debba tener conto soltanto quando si possa ritenere che grazie ad esse l'errore sul precetto avrebbe potuto non essere commesso.

L'inevitabilità dell'errore che in base al principio personalistico dovrebbe escludere la colpevolezza «non va misurata alla stregua [...] di parametri che valutino i dati influenti sulla conoscenza del precetto esclusivamente alla luce delle specifiche caratteristiche *personali* dell'agente», afferma la Corte, per poi affrettarsi a precisare che la «spersonalizzazione» di tale giudizio «va compensata [...] dall'esame di eventuali, particolari conoscenze ed “abilità” possedute dal singolo agente», poiché «queste ultime, consentendo all'autore del reato di cogliere i contenuti ed il significato determinativo della legge penale, escludono che l'ignoranza della legge penale vada qualificata come inevitabile»¹⁰⁶.

¹⁰⁴ Cass., Sez. Un., 25 gennaio 2005, n. 9163. Cfr. F. Mantovani, *Diritto penale*, cit., pp. 629 e 663 e, più recentemente M. Donini, *Responsabilità e pena da Kant a Nietzsche. La decostruzione del rimprovero*, in «Rivista italiana di diritto e procedura penale», LXIII, 2020, 4, pp. 1702-1703. Si veda tuttavia G.P. Fletcher, *Basic Concepts of Criminal Law*, Oxford University Press, New York-Oxford, 1998, trad. it. di M. Papa, *Grammatica del diritto penale*, il Mulino, Bologna, 2004, pp. 152-153: «Sebbene prevalente nella giurisprudenza del XX secolo, tanto nei paesi dell'Europa continentale che di *common law*, questo orientamento fu progressivamente abbandonato. Così, ad esempio, le corti tedesche cominciarono a richiedere che l'accusa provasse l'infondatezza anche delle eccezioni relative all'antigiuridicità e alla colpevolezza. Nell'area di *common law* le corti sono oggi fortemente divise in ordine alla possibilità di riversare sull'imputato l'onere di provare la legittima difesa o l'infermità mentale». Si veda anche F. Viganò, *Stato di necessità e conflitto di doveri*, cit., pp. 261-262 e 265.

¹⁰⁵ Corte cost., 24 luglio 2007, n. 322. Cfr., ad esempio, D. Pulitanò, *Diritto penale*, cit., p. 260 e G. Cocco, *La colpevolezza quale principio costituzionale (e il buon, vecchio libero arbitrio)*, cit., p. 6.

¹⁰⁶ Corte cost., 24 marzo 1988, n. 364. Cfr. A. Vallini, *Antiche e nuove tensioni tra colpevolezza e diritto penale artificiale*, cit., p. 287: «Paradossalmente, l'introduzione di una scusan-

In questo modo, d'altra parte, la Corte finisce in realtà per rafforzare la sensazione, sempre più diffusamente avvertita dalla stessa dottrina penalistica, che la colpevolezza *di per sé* non sia che «una sorta di illusorio “paravento” di modernità»¹⁰⁷ e che la natura «solo *convenzionale*, quasi *mistificatoria*» della sua affermazione non sia limitata a «certi ambiti»¹⁰⁸, ma costituisca la cifra stessa dell'ideologia penale che sostiene e giustifica l'esercizio della pretesa punitiva dello Stato, confermando ancora una volta, se mai ce ne fosse stato il bisogno, che «l'intero discorso che scienze sociali e scienze giuridiche alimentano intorno al sistema penale non fa che registrare lo scollamento tra principi e pratiche»¹⁰⁹.

Come ha constatato Ombretta Di Giovine in un suo scritto recente, la soluzione escogitata dai compatibilisti – i quali in genere sembrano accontentarsi «di argomentare la necessità di tenere distinti il piano neuro-filosofico da quello giuridico» per «difendere gli scopi che [...] persegue» il diritto – «si basa sull'idea che *un mondo in cui si finge che l'uomo sia libero è migliore di un mondo in cui lo si neghi* e per questa ragione è da preferire alle sue alternative»¹¹⁰. Il libero arbitrio tutt'al più sarebbe, dunque, una *finzione* «indispensabile» o comunque «estremamente utile per la prassi»¹¹¹ – come Vaihinger ipotizzava già agli inizi del secolo passato – e proprio questa, in fondo, è la critica che Caruso rivolge a Dennett in *Just Deserts*: tu ritieni che le persone dovrebbero essere trattate *come se* fossero «moralmente responsabili nel senso del merito»¹¹² non perché credi che effettivamente possano esserlo né perché consideri di per sé buono il «fatto che i trasgressori siano incolpati e puniti per le loro azioni»¹¹³, ma perché pensi *erroneamente* che senza un sistema basato sul merito staremmo tutti molto peggio¹¹⁴. Dennett nega ovviamente di aver proposto che

te viene ricollegata a presupposti del tutto oggettivi. Se poi qualche correttivo a tali criteri di massima si vuole porre, è un correttivo, si badi, *in malam partem*, rispetto al soggetto dotato di “particolari conoscenze ed abilità”. Non ci si preoccupa, invece, del soggetto con capacità, cultura, possibilità informative significativamente al di sotto della media, destinato ad essere inevitabilmente danneggiato da valutazioni incentrate su criteri generalizzanti».

¹⁰⁷ A. Vallini, *Antiche e nuove tensioni tra colpevolezza e diritto penale artificiale*, cit., p. 289.

¹⁰⁸ Ivi, p. 288.

¹⁰⁹ E. Resta, *La secolarizzazione della colpa. Note sugli autori del reato*, in L. Violante (a cura di), *Storia d'Italia. Annali 12. La criminalità*, cit., p. 127.

¹¹⁰ O. Di Giovine, *Delitto senza castigo?*, cit., pp. 257-258.

¹¹¹ H. Vaihinger, *La filosofia del “come se”*, cit., p. 54. Cfr. anche, ivi, p. 116.

¹¹² D.C. Dennett, G.D. Caruso, *A ognuno quel che si merita*, cit., p. 76. Cfr. anche, ivi, p. 60.

¹¹³ Ivi, p. 60.

¹¹⁴ Ivi, pp. 40-42, 76, 114 e ss.

«dovremmo, per così dire, *far finta* di ritenere le persone responsabili» *nel senso* in cui esse sono considerate tali dagli incompatibilisti e asserisce di avere affermato, invece, che «dovremmo ritenerle *realmente responsabili* nel senso non-intrinseco e non-assoluto» dei compatibilisti, «perché *sono realmente* responsabili in [quel] senso»¹¹⁵.

Ma poiché, come egli stesso ammette – ed anzi è «desideroso di “riconoscere”» – «la sua concezione è revisionista e cerca di difendere una nozione di libero arbitrio che è diversa da quella in cui crede la gente comune»¹¹⁶, data la posta in gioco e il portato normativo dei termini di cui si tratta (‘responsabilità’, ‘merito’, ‘rimprovero’, ‘colpevolezza’, ‘pena’, ecc.), è inevitabile che gli si chieda di farsi carico degli effetti persuasivi, se non mistificatori, delle sue ridefinizioni, ammesso che si tratti di ridefinizioni e non piuttosto di stipulazioni pure. E altrettanto vale, ovviamente, per le ridefinizioni o le stipulazioni cui ricorrono, spesso con gran disinvoltura e forse non sempre del tutto ingenuamente, giuristi, giudici e legislatori¹¹⁷.

Personalmente, per quel che vale, come Caruso fatico a comprendere come un mondo nel quale si infliggono deliberatamente ad alcuni dolore e sofferenza fingendo che se li siano “meritati” possa essere migliore d’uno in cui si crede che punire qualcuno che è “colpevole per finta” sia ipocrita, sleale e disonesto e finalmente si smette di farlo. La stessa Di Giovine, del resto, pur riconoscendo che «la libertà dell’uomo [...] a molti appare una finzione ancora necessaria»¹¹⁸, non esita a definire la soluzione dei compatibilisti «compromissoria e inappagante»¹¹⁹ e a ipotizzare che «il progredire dello Stato di diritto, prima o poi, ci porterà a fare i conti con il “fatto” che non siamo del tutto – e forse siamo davvero molto poco – liberi

¹¹⁵ Ivi, p. 109.

¹¹⁶ Ivi, p. 54.

¹¹⁷ Mi sono soffermato un po’ più diffusamente su questo problema in L. Milazzo, *Liberi tutti?*, cit., pp. 19 e ss. Qui vorrei limitarmi a osservare che convincono assai poco soluzioni come quella proposta da Hans Achenbach, il quale crede che, poiché «l’imputazione individuale di una condotta [...] può essere derivata e delimitata solamente a partire da una *interpretazione preventiva* degli elementi che fondano l’imputazione» e, dunque, prescindendo interamente dall’accertamento impossibile del fatto che l’autore avrebbe potuto agire diversamente, «si [dovrebbe] rinunciare del tutto al termine impreciso e confuso di “colpevolezza” („*Schuld*”)», senza che tuttavia questo comporti, per l’autore, che si debba rinunciare anche alla “pena” e, soprattutto, alla pratica sociale che la colpevolezza serve a legittimare (H. Achenbach, *Individuelle Zurechnung, Verantwortlichkeit, Schuld*, cit., pp. 137 e 140).

¹¹⁸ O. Di Giovine, *Delitto senza castigo?*, cit., p. 262.

¹¹⁹ Ivi, p. 258.

di scegliere se delinquere o meno»¹²⁰.

Quel che veramente conta, tuttavia, è che sia chiaro che cosa in pratica comporti considerare «l'alternativa *ex ante* tra possibilità di commettere e possibilità di omettere l'azione proibita che forma il presupposto della scelta [...] non già come ontologica ma come deontologica»¹²¹, ossia in sostanza assumere che dietro al principio di colpevolezza si nasconda nei nostri ordinamenti una norma che prescrive di trattare *come se* avessero avuto la possibilità di determinarsi ad agire lecitamente *alcuni*, perché *altri*, che condividono con loro alcune caratteristiche che l'ordinamento rende rilevanti, in genere lo fanno. Significa, nella migliore delle ipotesi, che «in base alle nostre regole giuridiche possiamo dichiarare colpevoli degli agenti che forse non lo sono affatto»¹²² o, se si preferisce, «che il caso e la necessità [...] possono rivelarsi all'origine della responsabilità penale, tanto quanto il dolo e la colpa»¹²³. Una volta che questo sia chiaro, ciascuno potrà fare le sue valutazioni.

¹²⁰ *Ibid.*

¹²¹ L. Ferrajoli, *Diritto e ragione. Teoria del garantismo penale*, Laterza, Roma-Bari, 2002 (7^a ed.), p. 504.

¹²² G. Stratenwerth, *Il concetto di colpevolezza nella scienza penalistica tedesca*, cit., p. 230.

¹²³ M. Donini, *Responsabilità e pena da Kant a Nietzsche*, cit., p. 1739. Mi permetto di rinviare anche, al riguardo, a L. Milazzo, *Sorte penale? La colpevolezza tra cause e caso*, in «Diritto & Questioni pubbliche», 14, 2014, pp. 794-858; Id., *Il libero arbitrio e i paradossi della colpevolezza*, in M. Lalatta Costerbosa (a cura di), *Lo spazio della responsabilità. Approdi e limiti delle neuroscienze*, il Mulino, Bologna, 2015, pp. 211-226.

*Il ruolo delle emozioni nel giudizio penale
tra “oggetto della valutazione” e “valutazione dell’oggetto”*

DANIELE PIVA*

Se, mutuando dal titolo di questo ciclo di seminari, si vuole andare “dentro” al giudizio penale il ruolo delle emozioni – come si è tentato di argomentare più ampiamente in un recente studio monografico cui non può che rinviarsi¹ – può apprezzarsi nella duplice prospettiva dell’imputato e del giudice potendo incidere, da un lato, sul coefficiente di rimproverabilità del comportamento (*oggetto della valutazione*) e, dall’altro, sui criteri della decisione (*valutazione dell’oggetto*): in altri termini, sul *cosa* e sul *come* giudicare.

Il primo terreno è senz’altro più dibattuto in ambito penalistico in quanto chiama in causa il giudizio di imputabilità e di colpevolezza nella loro dimensione costituzionale (art. 27 Cost.), il secondo vede invece la sua genesi nell’ambito di studi a carattere prevalentemente filosofico volti a esaltare il ruolo della compassione quale elemento costitutivo della discussione razionale a fondamento della sentenza². Ma entrambi, al fondo, richiamano l’esigenza di applicare la pena soltanto a chi, fisiologicamente influenzato da impulsi psicologici, la “merita” davvero e ne possa beneficiare in termini di riabilitazione, riportando al centro di tutto l’uomo (e non il reo), nella sua irripetibile unicità e per quello che è (e non per quello che *dovrebbe essere*), fatto di ragione e volontà ma anche di carattere, emozioni, sentimenti e incoscienza.

Partendo dunque dalla prospettiva del giudice, si sostiene che ad ostacolare la corretta formazione del suo libero convincimento sia l’indifferenza o, peggio, il disgusto e non già l’immedesimazione, dovendosi piuttosto valorizzare condizioni e motivazioni del reo nell’esercizio di una discrezionalità dotata, essa stessa, di appropriata emotività e capace di comprendere

* Dipartimento di Giurisprudenza, Università Roma Tre.

¹ D. Piva, *Le componenti impulsive della condotta: tra imputabilità, (pre)colpevolezza e pena*, Napoli, 2020, *passim*.

² Cfr., specialmente, M. Nussbaum, *L’intelligenza delle emozioni*, Bologna, Il Mulino, 2009, pp. 525ss.; Ead., *Nascondere l’umanità. Il disgusto, la vergogna, la legge*, Torino, Carocci, 2007, cit., pp. 70 ss.

le ragioni del crimine, per guardare all'imputato non come un nemico cui imporre insostenibili tensioni psichiche bensì come «persona fallibile e spesso indotta alla devianza da condizioni estranee [...] di disagio e deprivazione, o addirittura come vittima dell'etichettamento che gli affibbia lo stesso sistema sociale e punitivo»³. L'emozione diventa così strumento di *immaginazione* per compiere variazioni di prospettiva necessarie a rivivere il relativo sentire al momento dell'attivazione di un determinato impulso: in sintonia col principio noto in ambito filosofico come legge di Hume, nella misura in cui afferma l'impossibilità di derivare logicamente conclusioni valutative o normative sulla base di sole premesse fattuali ovvero con quello che gli psicologi statunitensi da tempo chiamano *heuristic affect* per indicare il contributo delle emozioni al processo decisionale⁴.

Alla compassione si riconosce un ruolo persino epistemico perché in grado di superare i pregiudizi, attirare l'attenzione su dettagli significativi, stimolare una riflessione critica e contribuire in modo efficace a ricercare e motivare le ragioni che giustificano una determinata condizione psicologica anziché farsi fatalmente condizionare dall'efferatezza di un atto o dall'apparente mostruosità del suo autore, se non anche dai suoi riflessi mediatici, nell'ambito di una rappresentazione parziale e selettiva ad alta densità assiologica.

Con riguardo al criterio di giudizio, non si tratta semplicemente di immedesimarsi negli impulsi altrui quanto di valutare il comportamento che ne sia scaturito in termini di appropriatezza, intesa come giustificabilità razionale. Il giudice sarebbe cioè tenuto ad adottare un approccio per *simpatia* o *antipatia* e non solo per *empatia*, dovendo sì rispecchiarsi nell'altrui stato psicologico ma previa una comprensione completa dei fatti che gli consenta di «penetrare le cause degli eventi»⁵: la differenza starebbe nel fatto di non dover soltanto percepire in modo esperienziale gli altrui impulsi psicologici ma di esprimere anche una valutazione critica sulla condotta, misurandone *merito* o *demerito* in modo imparziale. Detto altrimenti, empatia vuol

³ Così, per tutti, O. Di Giovine, *Un diritto penale empatico? Diritto penale, bioetica e neuroetica*, Torino, Giappichelli, 2009, pp. 162 s.

⁴ Cfr., tra i primi, M. L. Finucane, A. Ahakami, P. Slovic, S. M. Johnson, *The Affect Heuristic in Judgement of Risks and Benefits*, in «Journal of Behavioral Decision Making», 13, 2000, pp. 1-17; P. Slovic, M. L. Finucane, E. Peters, D. G. MacGregor, *Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality*, in «Risk Analysis», 24, 2004, pp. 21-36.

⁵ Il rinvio è a A. Cadoppi, *Simpatia, antipatia e diritto penale*, in O. Di Giovine (a cura di) *Diritto penale e neuroetica. Atti del Convegno 21-22 maggio 2012 Università degli studi di Foggia*, Padova, CEDAM, 2013, p. 243.

dire *partecipazione-mediante-distanziamento* del giudice⁶ mentre *simpatia* e *antipatia* presuppongono un *quid pluris* consistente nella condivisione mediante i sistemi riflessivi della conoscenza⁷: così, ad esempio, un episodio di adirata violenza suscita un sentimento di sdegno e riprovazione comunque diverso da quello derivante dall’effettiva conoscenza della provocazione a causa dell’ira o delle origini e dei motivi dell’azione.

In tal senso, la rilevanza delle componenti impulsive verrebbe a dipendere dal tasso di giustificabilità razionale del comportamento che ne sia derivato cui, a sua volta, corrisponde il risultato di una riflessione in termini di simpatia o antipatia (diretta, quanto ai sentimenti e ai motivi che lo sorreggono e specularmente indiretta quanto al risentimento di chi ne abbia subito le conseguenze).

Tale condivisione non è un’esperienza libera e solitaria del giudice ma passa attraverso un raffronto tra caso concreto e caso paradigmatico, quale situazione rappresentativa dello stereotipo di una determinata reazione in contesti relazionali, secondo la tecnica del *reasonable firmness test* mutuata dalla tradizione di *common law*⁸. Ed è qui che, in una logica di integrazione e non di esclusione, *simpatia* e *antipatia* vengono incanalate nel sistema delle regole che governano l’imputazione e delle massime di esperienza, impedendo che il giudizio si risolva nell’individuazione di tipi d’autore o dia sfogo a intuizioni o sensazioni arbitrarie, con inevitabile compromissione del principio di uguaglianza: nel tentativo, ove possibile, di convertire l’impulsivo da puro fattore di irrazionalità a criterio di misura relativo a interi ambiti di comportamento.

⁶ O. Di Giovine, *Una lettura evoluzionistica del diritto penale. A proposito delle emozioni*, in *Diritto penale e neuroetica. Atti del Convegno 21-22 maggio 2012 Università degli studi di Foggia*, cit., p. 338.

⁷ La differenza si coglie anche dal punto di vista etimologico: dal greco *εμπάθεια* (da *εμ-*, “dentro”, e *πάθος*, “sentimento”) significa comprendere l’altrui stato d’animo, come “mettersi nei panni dell’altro” mentre *συμπάθεια* (da *συν* + *πάσχω* = *συμπάσχω*, letteralmente “patire con”) e *ἀντιπάθεια*, (da *ἀντι-* “contro” e *πάθος*, “sentimento”) presuppongono, oltre alla comprensione, rispettivamente una condivisione o una disapprovazione.

⁸ Cfr., limitandosi alla manualistica, J. Horder, *Ashworth’s Principles of Criminal Law*, Oxford, Oxford University Press, 2019, pp. 129ss., pp. 142-144, pp. 228ss., pp. 239ss.; J. Loveless, M. Allen, C. Derry, *Complete Criminal Law*, Oxford, Oxford University Press, 2018, spec. pp. 385ss., pp. 426ss. e pp. 438ss.; M. Lippman, *Essential Criminal Law*, Thousand Oaks, Sage, 2016, pp. 128s.; F. Schmalleger, D. Hall, *Criminal Law Today*, Boston, Pearson, 2016, p. 155; N. Padfield, *Criminal Law*, Oxford, Oxford University Press, 2016, pp. 124ss.; G.E. Dix, *Criminal Law*, St Paul, 2016, spec. pp. 104, 110 e 123; J. D. Ohlin, *Criminal Law: Doctrine, Application, and Practice*, New York, Wolters Kluwer Law & Business, 2016, p. 736; R. Card, *Card, Cross & Jones: Criminal Law*, Oxford, Oxford University Press, 2012, pp. 704-707 e pp. 720-722.

Il giudizio comporta, inoltre, la combinazione tra il sapere dell'esperto e quello del giudice: il primo chiamato a stabilire tipologia e grado di incidenza dell'impulso sulle effettive capacità psicologiche di orientare la condotta criminosa al momento dell'agire e il secondo tenuto come *peritus peritorum* ad esprimere, nei limiti segnati dalle risultanze delle indagini empiriche e previa verifica del rispettivo rigore metodologico e argomentativo, la valutazione sul fatto nel suo complesso e la conseguente sua qualificazione giuridica in rapporto a valori, principi e istituti rilevanti. La prospettiva è quella di un dialogo nell'ambito del quale il giudice, lungi dall'appiattirsi su esiti peritali, non si sottragga ai vincoli costituzionali che governano l'imputazione, rientrando tra i suoi compiti non solo quello di fornire una spiegazione della realtà ma anche e soprattutto quello di garantire che in concreto l'applicazione della pena risulti giustificata. A lui spetta, in altri termini, il ruolo di "custode del metodo" (*gatekeeper*) col poterdovere di guidare l'operato dell'esperto fornendogli le nozioni giuridiche necessarie per l'individuazione dell'oggetto dell'indagine, nonché di controllare il rigore metodologico e argomentativo del procedimento dallo stesso utilizzato, per decostruirne le affermazioni ove dovessero emergere ragionevoli dubbi sulla loro validità, ma nello stesso tempo adeguarvisi se correttamente costruite, sulla base di un ragionamento prevalentemente induttivo condotto secondo massime di esperienza.

Da rimuovere, in quest'ottica *multilevel*, il limite del ricorso alla perizia psicologica (art. 220 cpv. c.p.p.) oggi ammessa solo in quanto volta ad accertare qualità psichiche dipendenti da cause patologiche che, viceversa, potrebbe garantire il ricorso alle migliori acquisizioni delle scienze empiriche in grado di esplorare i meccanismi della coscienza emozionale evitando giudizi basati su mere intuizioni ovvero un *non liquet* supportato da implicite presunzioni.

Venendo alla posizione dell'imputato, è noto come da tempo, per quanto timidamente, la giurisprudenza abbia dimostrato di allargare il giudizio di imputabilità rispetto al quadro delle rigide classificazioni cliniche dell'infermità mentale, anche al costo di forzature sistematiche e di un ampliamento della propria sfera di discrezionalità, facendo leva sulla prova di un rapporto diretto, eziologico e motivazionale, tra il reato e lo *status* dell'agente per stabilirne, ai sensi dell'art. 85 c.p., la capacità o meno di intendere e di volere in rapporto al singolo fatto commesso *hic et nunc*: in modo da escludere la pena in corrispondenza di «una situazione di assetto psichico incontrollabile e ingestibile [...] che, incolpevolmente, rende l'agente incapace di esercitare il dovuto controllo dei propri atti, di conseguentemente indirizzarli, di percepire il disvalore sociale del fatto, di

autonomamente, liberamente autodeterminarsi»⁹.

In questa categoria potrebbero allora rientrare le emozioni del reo quantomeno nei casi *border line* connessi, ad esempio, a scatti d'ira, gelosia delirante o crisi di panico tali da innescare reazioni a corto circuito in cui comunque si manifesti improvvisamente un impulso irresistibile (*raptus*) talvolta frutto di una lunga, pregressa ed elaborata situazione di *stress* (si pensi ai cd. delitti d'impeto) assimilabile a nevrosi o psicopatia. Se, infatti, ogni condotta trova origine in uno stato emotivo che risale agli istinti fondamentali della vita, ai sentimenti e al carattere della persona, la capacità di volere può considerarsi proprio come la possibilità di convertire o meno l'impulso in azione, venendo a mancare ogniqualvolta esso rimanga nello stato dell'incoscienza o sia di consistenza e gravità tale da eliminare l'effettiva volontà di contrapporre freni inibitori a fronte della consapevolezza di commettere un fatto ingiusto o antisociale. D'altronde, nella pratica psichiatrica e giudiziaria il confine tra iperemotività fisiologica e morbosa sembra ormai disperso nell'analisi della particolare intensità delle reazioni, della mancanza di una causa efficiente o del prolungamento oltre misura dei relativi effetti: sia per l'emotività diffusa, permanente e riscontrabile nella irritabilità dei collerici o isterici sia in quella sistematizzata rilevabile solo in determinate situazioni di spiccata paura degenerata in realtà illusoria prevaricante. Inoltre, gli studi della psicoanalisi sul debilitamento nervoso dimostrano che compulsione (come atto che si impone al soggetto su base conflittuale) e impulsione (come scarica di aggressività non mascherata) costituiscono entrambi espressione di emozioni che, non trovando una strada percorribile secondo le comuni e accettate regole della convivenza sociale, possono parimenti estrinsecarsi in condotte criminose i cui processi

⁹ Cfr. Cass., Sez. Un., 25.01.2005, n. 9163 (cd. sentenza Raso), in *Cass. pen.*, 2005, 1851 ss., con nota di Fidelbo, *Le Sezioni unite riconoscono rilevanza ai disturbi della personalità sulla quale, nella vasta letteratura di commento, v. M.T. Collica, Anche i "disturbi della personalità" sono infermità mentale*, in «Rivista italiana di diritto e procedura penale», 2005, pp. 420 ss. e Ead., *Vizio di mente: nozione, accertamento e prospettive*, cit., p. 108 ss.; M. Bertolino, *L'infermità mentale al vaglio delle Sezioni Unite*, in «Diritto penale e processo», 2005, p. 837ss.; Ead., *Empiria e normatività nel giudizio di imputabilità per infermità di mente*, in «Legislazione penale», 2006, pp. 212 ss.; D. Pulitanò, *La disciplina dell'imputabilità fra diritto e scienza*, in «Legislazione penale», pp. 248ss.; G. Fiandaca, *L'imputabilità nella interazione tra epistemologia scientifica ed epistemologia giudiziaria*, in «Legislazione penale», pp. 257 ss.; A. Manna, *L'imputabilità tra prevenzione generale e principio di colpevolezza*, in «Legislazione penale», pp. 220-247; I. Merzagora Betsos, *I nomi e le cose*, in «Rivista italiana di medicina legale», 2005, pp. 373 ss.; F. Centonze, *L'imputabilità, il vizio di mente e i disturbi della personalità*, in «Rivista italiana di diritto e procedura penale», 2005, p. 247; V. Militello, *Imputabilità, infermità di mente e disturbi della personalità nella evoluzione giurisprudenziale*, in «Giurisprudenza Penale», 2005, pp. 1601 ss.

incoscienti sono simili a quelli dell'affezione nevrotica¹⁰.

In tal senso si palesa l'illegittimità di una disposizione come quella dell'art. 90 c.p. che, nel condannare all'irrilevanza *tout court* gli stati emotivi e passionali (ai quali si *sa* resistere perché si *deve* saper resistere), costituisce norma di ottuso sbarramento priva di basi empiriche e comunque fondata sul riconoscimento alla pena di una funzione esclusivamente general-preventiva, se non addirittura su un giudizio di natura etica o pedagogica. Una norma, oltretutto, il cui effetto di "chiusura-rifiuto" deriva la sua forza dalla stessa natura tautologica della formulazione dalla quale non è in alcun modo possibile desumere, al di là di mere clausole di stile, cosa debba intendersi per emozioni o passioni e se tali concetti debbano esprimere il medesimo significato in riferimento ai diversi presupposti della responsabilità, attesa la mancanza di una definizione condivisa anche sul piano psicologico o filosofico e l'ulteriore limite probatorio imposto dall'art. 220 cpv. c.p.p. in relazione alla perizia su personalità, carattere e qualità psichiche indipendenti da cause patologiche. L'ennesimo tributo pagato sull'altare della finzione di un uomo incondizionatamente "libero" perché astratto da quello che realmente è ma semmai considerato per quello che *non può non essere*.

Ma è soprattutto sul piano della colpevolezza che le emozioni possono produrre l'effetto di "scusare" chi non sia riuscito a contenerle agendo altrimenti laddove non tanto manchi un potere psicologico di resistervi ma non possa più richiedersi - sulla base di una valutazione di carattere normativo che tenga conto della personalità e del carattere dell'agente e di tutte le condizioni empiricamente verificabili attinenti al relativo procedimento motivazionale nella situazione concreta - l'efficace attivazione dei propri freni inibitori. In tutti questi casi l'effetto dell'impulso emotivo partecipa di un insieme di circostanze anomale concomitanti (interne o esterne) che, nella valutazione legislativa, influiscono in modo irresistibile sulla volontà o sulle capacità psichiche, in guisa da non poter esigere da parte del soggetto un comportamento doveroso. A determinate condizioni non volontariamente causate dall'agente, l'ordinamento non punisce la condotta in quanto derivante da un processo motivazionale anomalo che, pur non pregiudicando il requisito della capacità di intendere e di volere, costituisce una reazione ragionevole a emozioni rispetto alle quali la resistenza non sia concretamente esigibile.

Esempi rinvenibili nel codice penale sono quelli della provocazione (art. 599 c.p.) in cui lo stato d'ira derivi da fatto ingiusto altrui, della costrizione

¹⁰ Cfr. B. Callieri, A. Castellani, G. De Vicentiis, *Trattato di psicopatologia forense*, Roma, Il Pensiero Scientifico, 1973, 605 ss.

(artt. 52, 54, 242 cpv. e 384 c.p.) per difesa o salvezza e, ancor più chiaramente dopo gli interventi della L. 36/2019, dell'eccesso di reazione per turbamento derivante da situazioni di pericolo (art. 55 cpv. c.p.) nelle quali risultino alterate la percezione della realtà o le capacità di orientare la propria condotta¹¹: a emozioni di questo tipo sono connessi, infatti, istinti di sopravvivenza e protezione (per sé o per altri) che costituiscono essenza primordiale dell'essere umano e ne influenzano, già a livello inconscio, tutte le possibili modalità di condotta attive (azioni o movimenti) o passive (omissioni o inerzia) sino a compromettere temporaneamente le funzioni di controllo a mo' di *blackout* potendo persino incidere sulla coscienza e volontà della condotta ai sensi dell'art. 42 c.p.

Per delimitarne l'effetto scusante senza eccedere in generalizzati indulgenzialismi potrebbe farsi riferimento ad alcuni indicatori di resistenza (riguardanti origine, natura e contenuto dell'impulso, tipologia, circostanze o effetti della reazione ovvero condizioni e/o alle qualità specifiche dell'agente) logicamente coordinati in modo da esprimere il maggior tasso di univocità o motivata concordanza possibile, nell'ambito di un giudizio di causalità psichica necessariamente indiziario (ma non arbitrario) per lo più fondato su massime di esperienza e valutazioni espresse in termini di ragionevole certezza.

L'importante sarà, ancora una volta, farsi guidare dal principio di colpevolezza (art. 27 Cost.), cui occorre restituire "carne e sangue", senza incappare in diffuso sentimentalismo ma nemmeno facendo finta di non vedere quello che non può non caratterizzare i procedimenti motivazionali a base della condotta umana, al più ricorrendo all'*escamotage* della salomonica concessione di attenuanti generiche ai sensi dell'art. 62-bis c.p.

Soltanto con un'adeguata considerazione degli effetti della dimensione emotiva, anziché rimanere ancorata a criteri astratti o persino evanescenti ovvero imbrigliata entro asfissianti logiche di puro normativismo, la responsabilità dell'imputato può assumere una dimensione autenticamente soggettiva, potendosi giustificare di fronte al singolo il rimprovero per un fatto che diviene espressione di una sua libertà del volere e non solo avvenimento a lui riconducibile in senso puramente materialistico: diversamente, la colpevolezza smarrirebbe non solo la sua funzione di garanzia ma anche e soprattutto quella educativa, lasciandosi contagiare da istanze prettamente repressive in spregio dei sottesi valori costituzionali.

¹¹ Per un'applicazione della scusante dell'eccesso per turbamento nella difesa legittima domiciliare cfr., volendo, Cass. Sez. I, III, 10 dicembre 2019 (ud. 10 ottobre 2019), n. 49883, in «Diritto penale e processo», 5/2020, 656 ss., con commento di D. Piva, *Oggettivo e soggettivo nell'eccesso di difesa per "grave turbamento"*.

Scatole nere, pregiudizi e discrezionalità

Tre questioni etiche su intelligenza artificiale e diritto

SOFIA BONICALZI*

SOMMARIO: 1. Introduzione – 2. Il problema della scatola nera – 3. Pregiudizi umani e algoritmici – 4. Giustizia distributiva e retributiva fra discrezionalità e uniformità – 5. Conclusioni.

1. Introduzione

Nel contributo prenderò sinteticamente in esame tre questioni etiche fra loro intrecciate e a loro volta legate all'utilizzo di sistemi di intelligenza artificiale (I.A.) nell'ambito del diritto: il problema della scatola nera (§ 2), il problema dei pregiudizi umani e algoritmici (§ 3) e la questione del rapporto fra uniformità e discrezionalità nell'amministrazione della giustizia distributiva e retributiva (§ 4).

Il problema della scatola nera (*black box*) concerne la mancanza di trasparenza e interpretabilità che caratterizza il processo decisionale algoritmico, soprattutto quando esso dipende da complesse forme di apprendimento profondo non supervisionato¹. Sebbene il problema sia trasversale ai vari settori in cui l'I.A. trova applicazione, esso appare particolarmente pressante nel caso delle sue implementazioni giuridiche, specialmente nell'ambito di sistemi non autocratici che valorizzino la condivisione del potere decisionale e la distribuzione delle informazioni rilevanti. Nel contesto della giustizia retributiva, ad esempio, e in una cornice di stampo democratico, gli imputati, già di per sé in una posizione vulnerabile, avrebbero infatti diritto a una spiegazione comprensibile dei meccanismi e degli esiti del processo decisionale, sia per prendervi parte attivamente sia per metterne eventualmente in discussione i risultati². Da

* Dipartimento di Filosofia, Comunicazione e Spettacolo, Università Roma Tre | Cognition, Values, Behaviour Lab (CVBE), Ludwig-Maximilians-Universität München.

¹ R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, in «22 Stanford Technology Law Review», vol. 242, no. 19-26, 2019.

² A. Završnik, *Criminal Justice, Artificial Intelligence Systems, and Human Rights*, in «ERA Forum», vol. 20, 2020, pp. 576-583.

questo punto di vista, la tensione attuale verso lo sviluppo di forme di *explainable* o *interpretable* I.A. appare, dunque, più che giustificata³.

Sarebbe, tuttavia, ingenuo rinunciare all'apporto che l'I.A. può offrire in termini di riduzione di costi e tempi collegati ai procedimenti giudiziari assumendo che il processo decisionale umano sia necessariamente più trasparente e razionale. I risultati delle scienze cognitive hanno infatti mostrato che questo è a sua volta affetto da molteplici zone d'ombra. In particolare, diverse forme di automatismi e pregiudizi cognitivi (*bias*) – agendo spesso al di fuori del controllo cosciente del soggetto, malgrado esperienza professionale ed educazione – potrebbero risultare altrettanto, se non ancora più, perniciosi per l'equità e la correttezza dei processi decisionali di stampo giuridico, impattando negativamente anche sulla fiducia che in essi ripongono i non addetti ai lavori⁴. Non è un caso, infatti, che l'I.A. sia spesso pubblicizzata come in grado non solo di imitare l'intelligenza umana, ma come capace di superarla, non essendo affetta dalle nostre più comuni fragilità cognitive (si veda, ad esempio, <https://www.equivant.com/>)⁵. Affidarsi all'I.A. garantirebbe, in linea di principio, una maggior uniformità dei processi decisionali e dei risultati corrispondenti, limando quegli eccessi di discrezionalità che sono tipici dell'intelligenza umana. Al momento l'obiettivo non è stato raggiunto, in quanto ci sono evidenze di come le applicazioni dell'I.A. non siano a loro volta immuni da pregiudizi (*bias* algoritmici) dovuti, ad esempio, al tipo di dati con cui gli algoritmi vengono addestrati o alla scarsa rappresentatività dei dati stessi. Senza contare i dubbi sollevati a proposito dell'effettiva superiorità dell'I.A. in termini di accuratezza predittiva⁶.

Tuttavia, anche laddove fosse possibile superare il problema della scatola nera e dei *bias* algoritmici – e sorvolando su altre importanti questioni etico-giuridiche, come la mancanza di una chiara *accountability* nel caso di errori commessi da *electronic persons*⁷ –, non è detto che rinunciare alla flessibilità e alla discrezionalità proprie del pensiero umano, e che

³ A. Rai, *Explainable AI: From Black Box to Glass Box*, in «Journal of the Academy of Marketing Science», vol. 48, 2020, pp. 137-141.

⁴ Cfr. J.A. Bargh, T.L. Chartrand, *The Unbearable Automaticity of Being*, in «American Psychologist», vol. 54, 1999, pp. 462-479; G. Edmond, K.A. Martire, *Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making*, in «The Modern Law Review», vol. 82, no. 4, 2019, pp. 633-664.

⁵ <<http://www.equivant.com>> (ultimo accesso 05.02.2023)

⁶ Cfr. J. Dressel, H. Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, in «Science Advances», vol. 4, no. 1, 2018.

⁷ Cfr. J.J. Bryson, M.E. Diamantis, T.D. Grant, *Of, for, and by the People: The Legal Lacuna of Synthetic Persons*, in «Artificial Intelligence and Law», 25, 2017, pp. 273-291.

si estendono alla capacità di adattare le regole alle circostanze in essere, abbia unicamente conseguenze positive, favorendo una maggior equità procedurale e sostanziale. Flessibilità e discrezionalità *buone* – difese anche da un recente report del Consiglio d'Europa sull'impatto che i processi decisionali algoritmici possono avere sui diritti umani⁸ – non sconfinano nell'arbitrio ma, vincolate dalle leggi stesse, sono, anzi, inerenti alla necessità di applicare regole e *standard* a contesti specifici. Esse hanno a che fare con la capacità di deliberare in modo diverso a partire da un medesimo set di premesse e di rispondere alla soggettività di chi è giudicato⁹. Se si vuole salvaguardarle occorre, allo stato attuale, bilanciare adeguatamente il contributo dell'A.I. e dell'intelligenza umana.

2. *Il problema della scatola nera*

Negli ultimi anni, a livello internazionale, si è assistito a una progressiva allocazione delle decisioni negli ambiti della giustizia distributiva (riguardante la distribuzione delle risorse a disposizione) e retributiva (riguardante la punizione dei trasgressori) all'I.A. Nel campo della giustizia retributiva, l'I.A., il cui uso è di solito limitato alla fase pre-processuale e post-processuale, può svolgere un ruolo, ad esempio, nell'individuazione dei (potenziali) criminali o nell'assegnazione dei casi a giudici specifici. Un caso interessante e abbastanza noto riguarda l'ideazione di *bots* o *avatar* per l'individuazione di criminali in rete – esempio ne è *Sweetie*, un *avatar* con le fattezze di una bambina filippina, creato dall'ONG *Terre des Hommes* per rintracciare pedofili *online*¹⁰. Nell'ambito della giustizia distributiva, l'I.A. è invece coinvolta, ad esempio, nelle decisioni relative all'assegnazione degli alloggi sociali, all'accesso e all'organizzazione dell'assistenza sanitaria o alle promozioni in ambito lavorativo¹¹.

⁸ Cfr. B. Wagner, *Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights*, Council of Europe, 2018, <<https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights>> (ultimo accesso 05.02.2023).

⁹ Cfr. R. Binns, *Human Judgement in Algorithmic Loops; Individual Justice and Automated Decision-Making*, in «Regulation and Governance», vol. 16, no. 1, 2022, pp. 197-211; M. Hildebrandt, *Algorithmic Regulation and the Rule of Law*, in «Philosophical Transactions of the Royal Society», vol. 376, no. 2128, 2018; R. Jorgensen, *Algorithms and the Individual in Criminal Law*, in «Canadian Journal of Philosophy», vol. 52, no. 1, 2022, pp. 61-77.

¹⁰ Cfr. A. Završnik, *Criminal Justice, Artificial Intelligence Systems, and Human Rights*, cit.

¹¹ Cfr. R. Binns, *Human Judgement in Algorithmic Loops; Individual Justice and Automated Decision-Making*, cit.; R.R. Jorgensen, *Algorithms and the Individual in Criminal Law*, cit.; A. Rai, *Explainable AI: From Black Box to Glass Box*, cit.

È importante sottolineare come, rispetto agli strumenti di supporto precedentemente esistenti, con “intelligenza artificiale” capace di agire come *primary decision-maker* si definisce un sistema che, a partire da un insieme di *input*, è in grado di costruire modelli dai quali ricava autonomamente raccomandazioni, previsioni o decisioni rispetto a un dato insieme di obiettivi (si veda il report *Unboxing Artificial Intelligence: 10 Steps to Protect Human Rights 2019*)¹². Questo rende l’I.A. profondamente diversa da precedenti ausili predittivi – come i più o meno sofisticati sistemi di *crime mapping*, ovvero strumenti che mettevano in relazione aree geografiche specifiche con i relativi tassi di criminalità –, la cui storia precede di molto l’avvento dell’I.A. e che necessitavano però dell’apporto umano per interpretare tendenze statistiche e formulare modelli¹³. Nel caso della moderna I.A., abbiamo a che fare, invece, con strumenti che sono in grado di prendere direttamente decisioni che toccano i soggetti giuridici¹⁴.

L’applicazione dell’I.A. a molti settori sociali sensibili – dalla sanità alla giurisprudenza – ha portato all’emergere di sfide etiche senza precedenti. Fra queste, particolarmente discusso è il cosiddetto “problema della scatola nera”. Come accennato, questo fa riferimento alla mancanza di trasparenza e interpretabilità del processo decisionale algoritmico, in particolare relativamente al modo in cui determinati *input* producono determinati *output*. In alcuni casi l’opacità si presenta come radicale, nel senso che gli stessi programmatori non sono in grado di decodificare come certi *output* siano stati prodotti. Tale opacità radicale, di natura essenzialmente tecnica, può essere dovuta a una molteplicità di cause, che agiscono da sole o in combinazione fra di loro. Ciò si verifica, per esempio, quando gli *output* derivano da una combinazione di variabili il cui numero e la cui complessità eccedono le nostre competenze cognitive *standard*, oppure quando entrano in gioco correlazioni complesse, modelli statistici o tecniche di apprendimento profondo irriducibili a spiegazioni di tipo argomentativo¹⁵.

In alternativa, l’incomprensibilità dei risultati può dipendere da fattori normativi, come nel caso in cui i meccanismi che regolano il processo

¹² <<https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights>> (ultimo accesso 05.02.2023).

¹³ J. Hunt, *From Crime Mapping to Crime Forecasting: The Evolution of Place-Based Policing*, in «National Institute of Justice Journal», last modified July 10, 2019, <<https://nij.ojp.gov/topics/articles/crime-mapping-crime-forecasting-evolution-place-based-policing>> (ultimo accesso 05.02.2023).

¹⁴ M. Hildebrandt, *Algorithmic Regulation and the Rule of Law*, cit.

¹⁵ R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, cit.

decisionale siano stati secretati. Relativamente ai procedimenti giudiziari, un esempio in tal senso è offerto dal caso, molto discusso, *Loomis v. Wisconsin*. Nel 2013, l'imputato, Eric L. Loomis, fu trovato colpevole di aver guidato un'auto utilizzata durante una sparatoria. Classificato come ad alto rischio di recidiva e condannato a sei anni di carcere e cinque di supervisione, Loomis si vide negare la possibilità di accedere a informazioni rilevanti circa le procedure e i metodi attraverso cui l'algoritmo predittivo COMPAS (acronimo di *Correctional Offender Management Profiling for Alternative Sanctions*) aveva emesso il suo rapporto di valutazione del rischio. Loomis presentò poi ricorso contro la sentenza, appellandosi al diritto a un giusto processo – tutelato, nel sistema statunitense, dal Quinto e dal Quattordicesimo Emendamento¹⁶. Tra le altre motivazioni portate a sostegno del ricorso vi era il fatto che l'opacità del processo decisionale algoritmico gli avrebbe impedito di valutarne l'accuratezza nonché la violazione del diritto a una sentenza personalizzata¹⁷.

Se *Loomis v. Wisconsin* rappresenta un caso molto noto, gli strumenti di predizione della recidiva basati su correlazioni statistiche che considerano variabili come età, storia lavorativa e precedenti criminali sono ampiamenti utilizzati nel sistema giuridico americano, per esempio per stabilire cauzioni o decidere a chi concedere la libertà vigilata¹⁸. Alcuni studi mostrerebbero un effettivo vantaggio predittivo per quanto riguarda la probabilità di identificare efficacemente i criminali a rischio di recidiva o che, se rilasciati, tenderanno a non presentarsi in fase processuale¹⁹. Tuttavia, quando le spiegazioni descrittive o tecniche circa le modalità con cui una certa decisione è stata presa risultano inaccessibili, diventa spesso difficile determinare se la corrispondente decisione normativa sia corretta. Questo significa che la mancanza di trasparenza nel processo decisionale potrebbe mascherare l'indifendibilità normativa di una certa conclusione, generando un ulteriore livello di opacità circa la legittimità degli esiti del

¹⁶ Cfr. K. Freeman, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, in «North Carolina Journal of Law & Technology», vol. 18, no. 75, 2016.

¹⁷ Cfr. J. Biddle, *On Predicting Recidivism: Epistemic Risk, Tradeoffs, and Values in Machine Learning*, in «Canadian Journal of Philosophy», 2020, pp. 1-21.

¹⁸ Cfr. A.M. Barry-Jester, B. Casselman, D. Goldstein, *Should Prison Sentences Be Based on Crimes That Haven't Been Committed Yet?*, in «FiveThirtyEight», last modified August 4, 2015.

¹⁹ Cfr. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, *Human Decisions and Machine Predictions*, in «The Quarterly Journal of Economics», vol. 133, n.1, 2017, pp. 237-93. Ma si veda anche J. Dressel, H. Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, cit.

processo decisionale²⁰.

Inoltre, la sottrazione di procedure decisionali complesse allo scrutinio pubblico comporta un rischio crescente di alienazione sia dei non addetti ai lavori sia degli esperti. In effetti, un assunto abbastanza comune (anche se potenzialmente fuorviante) è quello secondo cui le decisioni giudiziarie tradizionali rispecchiano in qualche misura il ragionamento umano e la sua logica quotidiana. In quanto tali, esse possono essere discusse o anche contestate attraverso strategie argomentative *standard*, alla portata sia degli esperti sia dei neofiti. Da un punto di vista sociologico, una radicale mancanza di comprensione, che colpisce soprattutto coloro che non hanno un'adeguata preparazione tecnica, implica un'accresciuta vulnerabilità nei confronti delle procedure giudiziarie e può generare squilibri di potere²¹. Ciononostante, una certa cautela nell'enfatizzare l'opacità tipica dei sistemi basati sull'I.A. resta doverosa. Infatti, anche i procedimenti giudiziari tradizionali possono nascondere delle zone d'ombra, per esempio quando ci si affida a testimoni anonimi o a informazioni che non vengono rivelate alla controparte. Tuttavia, nei procedimenti giudiziari tradizionali l'uso di questo tipo di prove o testimonianze viene solitamente limitato il più possibile e deve essere conforme a rigide norme, piuttosto che essere considerato una parte quasi inevitabile del processo²².

A fronte di queste difficoltà, negli ultimi anni si è posta una maggior attenzione a come rendere le procedure decisionali dell'I.A. maggiormente interpretabili o spiegabili. L'obiettivo è naturalmente quello di sviluppare metodi e processi che possano essere meglio compresi, e potenzialmente controllati, dagli esseri umani²³. Dal punto di vista tecnico, tuttavia, la richiesta di trasparenza non è facile da soddisfare, soprattutto perché il miglioramento delle funzionalità si accompagna spesso a un corrispondente aumento della complessità. Dal punto di vista teorico, questa richiesta deve essere poi maggiormente articolata, *in primis* specificando il livello di comprensibilità richiesto, determinando il tipo di contenuti che devono essere comunicati alle parti e stabilendo la quantità di informazioni che serve agli agenti per prendere decisioni effettivamente più informate senza

²⁰ Cfr. A. D. Selbst, S. Barocas, *The Intuitive Appeal of Explainable Machines*, in «Fordham Law Review», vol. 87, no. 3, 2018.

²¹ Cfr. R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, cit.

²² A. Završnik, *Criminal Justice, Artificial Intelligence Systems, and Human Rights*, cit.

²³ P. Hall, N. Gill, *An Introduction to Machine Learning Interpretability*, Sebastopol (CA), O'Reilly Media, Incorporated, 2019, (II ed.); P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, *Explainable AI: A Review of Machine Learning Interpretability Methods*, in «Entropy», vol. 23, no. 1, 2020, p. 18; A. Rai, *Explainable AI: From Black Box to Glass Box*, cit.

incorrere in eccessivi oneri cognitivi²⁴. È, tuttavia, importante notare che il tentativo di generare spiegazioni comprensibili può sfociare in resoconti intuitivi ma non realmente rappresentativi dei processi sottostanti che sfuggono, per loro stessa natura, al ragionamento proposizionale²⁵.

Paradossalmente, mentre l'imperscrutabilità di per sé può spingere al rifiuto dei giudizi formulati da sistemi basati sull'I.A.²⁶, si osservano anche fenomeni opposti, con la tendenza a riporre un'eccessiva fiducia nei processi decisionali algoritmici a scapito delle controparti umane, verso le quali si finisce per nutrire un crescente scetticismo. Può generarsi così una dinamica auto-rinforzante che progressivamente aumenta la pressione sociale a sostituire le competenze di agenti umani con quelle degli agenti artificiali, visti come più efficaci – seppur misteriosi – risolutori dei problemi sociali esistenti. Questa dinamica è di per sé pericolosa quando porta a un'eccessiva fiducia nei confronti delle macchine e al mancato esercizio di forme di monitoraggio, rinunciando a mantenere un *human-in-the loop*²⁷.

Da un punto di vista psicologico, questa tendenza si inserisce in un discorso più ampio su come il ragionamento umano sia influenzato dall'interazione con l'I.A. e su come gli esseri umani si trovino a riporre fiducia nelle macchine in situazioni di incertezza e vulnerabilità²⁸. Anche al di fuori del contesto legale, affidarsi all'I.A. e ad altri sistemi automatizzati di ausilio alle attività quotidiane e professionali è di per sé associato a una serie di modifiche della cognizione e delle conseguenti prestazioni. Fra queste è particolarmente rilevante la cosiddetta *degradation skill*, che fa riferimento alla progressiva perdita di abilità non più esercitate in quanto demandate a un agente esterno automatizzato. In ambito professionale,

²⁴ J. Biddle, *On Predicting Recidivism: Epistemic Risk, Tradeoffs, and Values in Machine Learning*, cit.

²⁵ C.E. Perez, *Deep Learning's Uncertainty Principle*. April 6, 2018. <<https://medium.com/intuitionmachine/deep-learning-uncertainty-principle-13f3ffdd15ce#:~:text=The%20uncertainty%20principle%20as%20applied,interpretable%20don%27t%20generalize%20well>> (ultimo accesso 05.02.2023); R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, cit.

²⁶ A. Rai, *Explainable AI: From Black Box to Glass Box*, cit.

²⁷ R. Binns, *Human Judgement in Algorithmic Loops; Individual Justice and Automated Decision-Making*, cit.

²⁸ A. Jacovi, A. Marasović, T. Miller, Y. Goldberg, *Formalizing Trust in Artificial Intelligence: Prerequisites, Causes and Goals of Human Trust in AI*, in «Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency», 2021, pp. 624-635; J. Zerilli, U. Bhatt, A. Weller, *How Transparency Modulates Trust in Artificial Intelligence*, in «Patterns», vol. 3, no. 4, 2022.

l'impatto dell'I.A. sulle competenze dipende comunque, in misura non indifferente, dal tipo di lavoro dell'utente, con un potenziale aumento delle competenze nei lavori altamente qualificati e un effetto opposto su quelli non qualificati²⁹. La *degradation skill* si accompagna poi, spesso, a fenomeni quali l'*automation complacency*, che consiste nell'affidarsi in modo acritico e passivo a tecnologie almeno apparentemente più competenti dell'utente, e l'*automation bias*, per cui si tende a preferire soluzioni automatizzate in quanto ritenute più affidabili, specialmente nel caso in cui queste differiscano dai risultati ottenuti dall'utente³⁰.

3. Pregiudizi umani e algoritmici

Se uno stretto rapporto con le nuove tecnologie promuove modifiche della cognizione, quali *degradation skill*, *automation complacency* e *automation bias*, una vasta letteratura nell'ambito delle scienze cognitive e sociali ha enfatizzato come le facoltà cognitive umane siano naturalmente soggette a limiti e distorsioni. In particolare, molte nostre decisioni apparentemente razionali sarebbero influenzate o in larga parte determinate da processi automatici o inconsci che eludono la riflessione metacognitiva³¹.

Anche i giudici professionisti non sono immuni da *bias* di vario tipo che possono manifestarsi ed esercitare la loro influenza in varie fasi del procedimento giudiziario, dalla raccolta delle prove all'applicazione delle leggi ai casi specifici³². A questo proposito, una distinzione abbastanza comune nella *common law* è quella fra *actual* e *apprehended bias*. I primi si presentano quando il giudice nutre pregiudizi che lo portano a non tenere in adeguata considerazione le evidenze raccolte. Questo in virtù, per esempio, di qualche interesse personale o professionale o dell'esposizione a informazioni contrastanti con un forte potere persuasivo. I secondi sono chiamati in causa quando il pubblico o gli imputati stessi appaiono

²⁹ J.R. Holm, E. Lorenz, *The Impact of Artificial Intelligence on Skills at Work in Denmark*, in «New Technology, Work and Employment», vol. 37, no. 1, 2022, pp. 79-101.

³⁰ K. Goddard, A. Roudsari, J.C. Wyatt, *Automation Bias: Empirical Results Assessing Influencing Factors*, in «International Journal of Medical Informatics», vol. 83, no. 5, 2014, pp. 368-75; R. Parasuraman, D.H. Manzey, *Complacency and Bias in Human Use of Automation: An Attentional Integration*, in «Hum Factors», vol. 52, no. 3, 2010, pp. 381-410.

³¹ Cfr. A. Bargh, J. A., T.L. Chartrand, *The Unbearable Automaticity of Being*, cit.

³² G. Edmond, K.A. Martire, *Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making*, cit.

ragionevolmente preoccupati circa il fatto che il giudice non sarà in grado di prendere decisioni imparziali³³.

Gli studi a disposizione mostrano come fattori psicologici, come l'umore del giudice o dell'impiegato, o caratteristiche solo apparentemente irrilevanti, come l'aspetto fisico degli imputati o di chi fa richiesta di un mutuo, possano contribuire a spiegare la variabilità nelle decisioni umane³⁴. Ciò malgrado la tendenza diffusa fra i professionisti del diritto a considerarsi straordinariamente resistenti ai pregiudizi, in virtù di un presunto *eccezionalismo* dovuto a formazione ed esperienza, e nonostante la capacità di mostrarsi apparentemente imparziali³⁵. Se la mancanza di trasparenza rimane un problema nell'interazione con l'I.A., è possibile, dunque, appellarsi ai risultati delle scienze cognitive per mostrare che anche il ragionamento umano può essere altrettanto opaco, con fenomeni di auto-inganno per cui l'agente resta ignaro delle ragioni reali per le quali ha preso una certa decisione e con potenziali implicazioni negative sull'equità e la correttezza del processo decisionale.

Un esempio significativo della discussione in tal senso è costituito dal dibattito sulla diffusione dei pregiudizi impliciti fra i professionisti in ambito legale. Com'è noto, i pregiudizi impliciti sono legati a forme di discriminazione involontaria e che riflette stereotipi sociali problematici, potenzialmente in grado di distorcere il giudizio esplicito o il comportamento che ne deriva³⁶. Alcuni studi mostrano che i giudici sono affetti da pregiudizi analoghi a quelli diffusi nella popolazione ma, se adeguatamente motivati, sono in grado di limitarne l'effetto³⁷, con la conseguente discussione circa l'appropriatezza delle contromisure per contrastare eventuali pratiche discriminatorie³⁸. Inoltre, nonostante l'abbondanza e la disponibilità di informazioni sull'esistenza dei pregiudizi impliciti, la probabilità che i giudici, anche se ben informati, prendano sul serio l'ipotesi per cui i trasgressori siano mossi da tali pregiudizi rimane modesta³⁹.

³³ Ivi.

³⁴ Cfr. J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, S. Mullainathan, *Human Decisions and Machine Prediction*, cit.

³⁵ G. Edmond, K. A. Martire, *Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making*, cit.

³⁶ J. Holroyd, *Implicit Bias, Awareness and Imperfect Cognitions*, in «Consciousness and Cognition», vol. 33, 2015, pp. 511-523.

³⁷ J.J. Rachlinski, S.L. Johnson, *Does Unconscious Racial Bias Affect Trial Judges*, in «Notre Dame Law Review», vol. 84, no. 3, 2009.

³⁸ J. Kang, M.R. Banaji, *Fair Measures: A Behavioral Realist Revision of "Affirmative Action"*, in «California Law Review», vol. 94, no. 4, 2006, pp. 1063-1118.

³⁹ J. E. Girvan, *On Using the Psychological Science of Implicit Bias to Advance Anti-Discrimination*

Inoltre, i dati a disposizione dimostrano che i giudici sono affetti da pregiudizi cognitivi comuni⁴⁰, tra cui l'*effetto ancoraggio*, in base a cui l'esposizione a informazioni numeriche arbitrarie influisce sull'entità delle sentenze comminate⁴¹, e l'*effetto aspettativa*, per cui le convinzioni sulla colpevolezza degli imputati vengono involontariamente trasmesse ai giurati attraverso forme di comunicazione non verbale, influenzando così il verdetto finale⁴². L'impatto di fattori contingenti, presumibilmente irrilevanti, nella decisione giudiziaria è stato evidenziato da un discusso studio che mette alla prova il detto popolare secondo cui la giustizia è "ciò che il giudice ha mangiato a colazione". Lo studio mostra che la percentuale di sentenze favorevoli pronunciate da alcuni giudici durante una tipica giornata lavorativa diminuisce progressivamente per essere poi ripristinata dopo una pausa. Gli autori dello studio suggeriscono che lo sforzo decisionale esaurisce progressivamente le funzioni esecutive e le risorse mentali dei giudici, progressivamente più propensi a non intervenire e ad accettare lo *status quo*, in questo caso respingendo la richiesta dell'imputato⁴³.

Nel contesto giuridico, i pregiudizi cognitivi sono particolarmente preoccupanti in virtù del fatto che l'integrità dei processi giudiziari ha, fra i suoi elementi cardine, l'indipendenza e l'imparzialità dei giudici e il fatto che le decisioni siano prese esclusivamente sulla base di prove ammissibili. A fronte di tali difficoltà, non sorprende che l'I.A. sia presentata come in grado di fornire un valido supporto ai processi giudiziari tradizionali, superando la preoccupante variabilità tipica del processo decisionale umano. Rispetto a quest'ultimo, l'I.A. si presenta come più stabile ed efficiente: l'accuratezza delle prestazioni non decade nel tempo e non è soggetta a influenze ambientali contingenti. Queste caratteristiche consentono di ridurre i tempi di lavoro e di garantire una maggior uniformità nelle prestazioni.

Un problema importante, tuttavia, è che non è affatto chiaro fino a che punto i sistemi di I.A. attualmente disponibili nel settore legale

Law, in «George Mason University Civil Rights Law Journal», vol. 26, no. 3, 2015.

⁴⁰ G. Edmond, K.A. Martire, *Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making*, cit.

⁴¹ B. Englich, T. Mussweiler, F. Strack, *Playing Dice with Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making*, in «Personality and Social Psychology Bulletin», vol. 32, no. 2, 2006, pp. 188-200.

⁴² R. Rosenthal, *Covert Communication in Laboratories, Classrooms, and the Truly Real World*, in «Current Directions in Psychological Science», vol. 12, no. 5, 2003, pp. 151-154.

⁴³ S. Danziger, J. Levav, L. Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, in «PNAS», vol. 108, no. 17, 2011, pp. 6889-6892.

possano prendere decisioni imparziali e prive di pregiudizi⁴⁴. Il problema del pregiudizio algoritmico (*algorithmic bias*), ora ufficialmente affrontato da alcune discusse iniziative legislative (si veda, ad esempio, il rapporto di Human Rights Watch, *How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers*⁴⁵, sulla proposta di legge europea sull'I.A.), si riferisce a errori sistematici nei sistemi informatici che produrrebbero, di conseguenza, risultati ingiusti a vantaggio di determinati gruppi sociali, finendo per rafforzare gli stereotipi esistenti⁴⁶. A questo proposito, un ampio dibattito è stato suscitato dall'*audit* condotto nel 2016 da ProPublica, un'organizzazione no-profit con sede a New York e dedita al giornalismo investigativo⁴⁷. Il report di ProPublica⁴⁸ si concentra sul già citato COMPAS – un sistema di valutazione del rischio di recidiva, implementato da Equivant, un'azienda produttrice di software per il mondo giuridico, e ampiamente utilizzato negli Stati Uniti – mostrando come questo fosse prevenuto nei confronti della popolazione afroamericana. Nonostante l'etnia non sia una delle variabili direttamente considerate dal *software*, le previsioni dell'algoritmo possono comunque condurre a discriminazioni razziali dal momento che esse includono altre variabili spesso collegate alla provenienza etnica⁴⁹. Secondo il report di Angwin e colleghi⁵⁰, per esempio, l'errore nello stabilire la probabilità di recidiva colpiva in modo diverso imputati bianchi e neri, di fatto sottostimando la probabilità di recidivare dei primi e sovrastimando la probabilità di recidivare dei secondi⁵¹.

⁴⁴ J. Dressel, H. Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, cit.

⁴⁵ <<https://www.hrw.org/news/2021/11/10/how-eus-flawed-artificial-intelligence-regulation-endangers-social-safety-net>> (ultimo accesso 05.02.2023).

⁴⁶ L. Belenguer, *AI Bias: Exploring Discriminatory Algorithmic Decision-Making Models and the Application of Possible Machine-Centric Solutions Adapted from the Pharmaceutical Industry*, in «AI and Ethics», vol. 2, 2022, pp. 771-787.

⁴⁷ J. Angwin, J. Larson, L. Kirchner, *Machine Bias*, in «ProPublica», May 23, 2016.

⁴⁸ Per una risposta, si veda W. Dieterich, C. Mendoza, T. Brennan, *COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity. Performance of the COMPAS Risk Scales in Broward County*, Northpointe Inc. Research Department, July 8, 2016. <https://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf> (ultimo accesso 05.02.2023).

⁴⁹ Si veda anche D. Pedreshi, S. Ruggieri, F. Turini, *Discrimination-Aware Data Mining*, in «Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining», 2008, pp. 560-568.

⁵⁰ J. Angwin, J. Larson, L. Kirchner, *Machine Bias*, cit.

⁵¹ J. Dressel, H. Farid, *The Accuracy, Fairness, and Limits of Predicting Recidivism*, cit.

4. Giustizia distributiva e retributiva fra discrezionalità e uniformità

Anche supponendo che il problema del pregiudizio algoritmico – al momento una delle sfide principali per gli usi giuridici dell’I.A. – possa essere risolto, non è chiaro se la neutralità e l’uniformità dei risultati che l’I.A. può raggiungere siano sempre positive. Com’è noto, una distinzione *standard* che appartiene alla tradizione della *common law* statunitense è quella fra *formalismo* e *realismo* giuridico. Secondo i formalisti, il processo decisionale giuridico consiste nell’applicazione meccanica e logica di regole e ragioni predefinite. Il diritto rappresenta “un sistema concettuale completo, formale e ordinato, in grado di fornire per ogni caso risposte univoche e corrette”⁵². Al contrario, i realisti ritengono che una serie di fattori psicologici, contestuali e culturali plasmino le decisioni degli arbitri. Questi deliberano, quindi, in base a ciò che sembra loro giusto in una determinata situazione⁵³. Se ciò è particolarmente vero nei paesi basati sulla *common law* – cui il presente contributo fa maggiormente riferimento –, il tema si pone anche nei sistemi che adottano forme di *civil law*, ogniquale volta dei principi generali debbano essere adattati, con una ben calibrata dose di discrezionalità, a circostanze specifiche. Del resto, il confronto fra il ruolo dei giudici nei sistemi di *common* e *civil law* suggerisce che uno degli aspetti ove vi è maggiore convergenza riguarda proprio la valorizzazione della competenza discrezionale. Anche nei paesi dove vige la *civil law*, i giudici finiscono spesso per svolgere una sorta di funzione legislativa, impegnandosi appunto nell’adattamento di principi generali a circostanze specifiche⁵⁴.

Il già citato studio di Danziger, Levav e Avnaim-Pesso⁵⁵, in cui le decisioni dei giudici dipendono da quanto tempo è passato dalla colazione, offre una versione caricaturale del realismo giuridico. Lo spirito del realismo è però forse espresso al meglio dal famoso detto di Oliver Wendell Holmes,

⁵² C. Faralli, *L’eredità del realismo giuridico americano*, in «Materiali per una storia della cultura giuridica», vol. 36, no. 1, 2006, pp. 119-26, cit. p. 119.

⁵³ B. Leiter, *American Legal Realism*, in W. Edmundson and M. Golding (eds.), *The Blackwell Guide to Philosophy of Law and Legal Theory*, Oxford, Blackwell, 2005, pp. 50-66; R.A. Posner, *Legal Formalism, Legal Realism, and the Interpretation of Statutes and the Constitution*, in «Case Western Reserve Law Review», vol. 37, no. 179, 1986.

⁵⁴ S.B. Yu, *The Role of the Judge in the Common Law and Civil Law Systems: The Cases of the United States and European Countries*, in «International Area Studies Review», vol. 2, no. 2, 1999, pp. 35-46.

⁵⁵ S. Danziger, J. Levav, L. Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, cit.

secondo cui “la vita della legge non è stata logica; è stata esperienza”⁵⁶. Nonostante si tratti di un commento descrittivo relativo a come i processi decisionali si sviluppino in ambito legale, esso acquisisce una forza normativa nella misura in cui l’equità nelle decisioni può essere raggiunta proprio attraverso l’esercizio di una ragionevole discrezionalità plasmata, appunto, dall’esperienza. Se l’imparzialità rimane, quindi, un principio cardine di un procedimento giudiziario corretto ed equo, è inevitabile che i giudici abbiano prospettive, valori ed esperienze personali – per quanto queste possano entrare in tensione con la legittima aspettativa che i diversi casi vengano affrontati in modo omogeneo e coerente o che i vari gruppi sociali non vengano avvantaggiati o svantaggiati in modo sproporzionato⁵⁷. Enfatizzare la presunta neutralità e uniformità prestazionale dell’I.A. suggerisce che flessibilità e discrezionalità, tipiche dell’intelligenza umana, debbano essere limitate il più possibile. In particolare, promettendo *performance* efficienti e un’imparzialità almeno apparente, l’I.A. favorisce la svolta verso una ‘giustizia codificata’, che privilegia la standardizzazione rispetto alla discrezionalità⁵⁸.

L’eccesso di discrezionalità rappresenta, naturalmente, un serio ostacolo al raggiungimento di decisioni eque, in particolare quando sconfinata nell’arbitrarietà o in forme di discriminazione. Tuttavia, flessibilità e discrezionalità possono al contrario svolgere un ruolo sociale positivo, laddove permettano di risolvere iniquità esistenti e a patto, naturalmente, che non danneggino gruppi sociali già svantaggiati: la necessità di astenersi dall’arrecare ulteriori svantaggi a coloro che sono già svantaggiati rappresenta, infatti, un principio etico di base che è assunto nella maggior parte delle teorie della giustizia⁵⁹.

A questo proposito, politologi e scienziati sociali hanno osservato come l’applicazione meccanica e standardizzata delle regole esistenti possa causare problemi soprattutto a chi è socialmente svantaggiato, potenzialmente aggravando le forme di discriminazione esistenti. Alcuni esempi significativi in tal senso, nell’ambito della giustizia retributiva e distributiva, sono

⁵⁶ O.W. Holmes, *The Common Law*. Mineola, NY, Dover Publications, 1881/1991.

⁵⁷ R. Binns, *Human Judgement in Algorithmic Loops; Individual Justice and Automated Decision-Making*, cit.; G. Edmond, K.A. Martire, *Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making*, cit.

⁵⁸ M. Hildebrandt, *Algorithmic Regulation and the Rule of Law*, cit.; R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, cit.

⁵⁹ J. Biddle, *On Predicting Recidivism: Epistemic Risk, Tradeoffs, and Values in Machine Learning*, cit.

discussi da Virginia Eubanks nel saggio *Automating Inequality*⁶⁰. Per quanto riguarda la giustizia retributiva, Eubanks cita un rapporto della *Leadership Conference on Civil and Human Rights* (2000) che fa il punto su diverse leggi sulle pene minime obbligatorie emanate negli Stati Uniti nei decenni precedenti e che limitano la discrezionalità dei giudici. Sulla base delle prove acquisite in termini di disparità razziale nell'ambito della giustizia penale, il rapporto afferma esplicitamente che le minoranze si trovano maggiormente sfavorite in un sistema di condanne obbligatorie rispetto a quanto accade loro in un sistema che favorisce la discrezionalità giudiziaria. Laddove i giudici sono privati dell'autorità necessaria per imporre sentenze giuste, le leggi e le linee guida sulle condanne obbligatorie e l'automatizzazione dei processi giudiziari finiscono per rafforzare le disparità esistenti. Questo esito appare paradossale, nella misura in cui una delle principali giustificazioni per l'automatizzazione della giustizia è proprio quella di ridurre gli squilibri nel trattamento dei diversi gruppi sociali.

Per quanto riguarda la giustizia distributiva, l'impatto sociale delle decisioni algoritmiche è evidente nel caso della definizione dei profili di rischio in vista dell'assegnazione di alloggi a senzatetto. Un caso emblematico è offerto dal programma di edilizia sociale *Home for Good*, implementato nel 2013 per aiutare i senzatetto nell'area degradata di Skid Row, a Los Angeles. Semplificando i metodi preesistenti, attraverso uno strumento di valutazione che raccoglie informazioni e classifica i senzatetto in base a un "punteggio di vulnerabilità", il programma avrebbe consentito di limitare l'impatto dei pregiudizi impliciti ed espliciti spesso diffusi fra gli impiegati.

Sebbene il programma si sia rivelato efficace nell'assistere svariate persone con una storia pregressa di instabilità abitativa, alcuni fra i soggetti intervistati riferivano di essersi visti negare automaticamente la possibilità di ricevere un alloggio: il sistema, di per sé opaco, non forniva dettagli chiari su come il punteggio di vulnerabilità fosse stato calcolato. Inoltre, raccogliendo informazioni sugli aspiranti candidati, il sistema finiva per tracciare e monitorare i senzatetto, di fatto esasperando le disparità esistenti: le persone il cui comportamento e stile di vita erano classificati dall'indice di vulnerabilità come rischiosi o illegali ottenevano una priorità alta. Questa avrebbe sì consentito un accesso più rapido alle abitazioni, ma con l'effetto negativo di esporli al tempo stesso a un controllo più severo e potenzialmente alla detenzione. In quanto tale il programma non

⁶⁰ V. Eubanks, *Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor*, New York, NY, St Martins Pr., 2018.

si limitava a collegare la domanda abitativa con le risorse disponibili, ma si configurava come uno strumento di sorveglianza sociale e controllo degli emarginati. Questo senza però al contempo offrire quell'attenzione individualizzata e quella capacità di adattare le regole alle circostanze che sono più tipiche delle forme tradizionali di assistenzialismo⁶¹. Del resto, che una delle potenzialità dell'I.A. sia quella di fornire strumenti regolatori della società nel suo insieme è testimoniato da casi eclatanti, come quello di Singapore, spesso descritta come un laboratorio per testare i limiti di una comunità dove l'ordine nazionale è mantenuto anche attraverso l'accentramento tecnologico e la sorveglianza di massa⁶².

Azzerando flessibilità e discrezionalità, l'I.A. può, quindi, obliterare il senso stesso della legge – un punto sollevato fin dalle prime analisi sul rapporto tra tecnologia e diritto⁶³: essa non si limita a influire sulle caratteristiche più superficiali dei meccanismi giudiziari (ad esempio, in termini di riduzione dei tempi e dei costi), ma incide indirettamente sui valori che modellano la cultura giuridica esistente, potenzialmente a scapito dell'equità procedurale e sostanziale. Se da un lato i sistemi di I.A. possono garantire buoni risultati per quanto concerne l'uniformità delle prestazioni, dall'altro una sfida specifica è rappresentata, dunque, dalla necessità di preservare una certa dose di discrezionalità nelle decisioni giudiziarie.

Che il senso di giustizia non sia riducibile all'applicazione meccanica di regole preesistenti è, del resto, ampiamente dimostrato dai risultati delle scienze cognitive. Nell'ambito della giustizia distributiva e retributiva, i giudizi relativi al giusto e allo sbagliato sono resi possibili dall'integrazione flessibile di molteplici processi psicologici, tra cui figurano fattori sia cognitivi sia emotivi – per quanto ogni distinzione netta fra cognizione ed emozione vada accantonata⁶⁴. Ciò è reso evidente, ad esempio, dai test di economia sperimentale basati sulla distribuzione delle risorse disponibili, *in*

⁶¹ Per una serie di esempi su come un uso attento della discrezionalità possa migliorare le condizioni delle fasce svantaggiate, si veda, invece, M. Kelly, *Theories of Justice and Street-Level Discretion*, in «Journal of Public Administration Research and Theory», vol. 4, 1994, pp. 119-140.

⁶² S. Harris, *The Social Laboratory*, in «Foreign Policy», last modified July 29, 2014. Harris 2014 <<https://foreignpolicy.com/2014/07/29/the-social-laboratory/>> (ultimo accesso 05.02.2023); A. Završnik, *Criminal Justice, Artificial Intelligence Systems, and Human Rights*, cit.

⁶³ Si veda ad esempio P. Leith, *Fundamental Errors in Legal Logic Programming*, in «The Computer Journal», vol. 29, 1986, pp. 545-552.

⁶⁴ H. Okon-Singer, T. Hendler, L. Pessoa, A.J. Shackman, *The Neurobiology of Emotion–Cognition Interactions: Fundamental Questions and Strategies for Future Research*, in «Frontiers Human Neuroscience», vol. 9, 2015.

primis nei risultati tipici del gioco dell'*ultimatum*. Nel gioco dell'*ultimatum*, uno dei partecipanti ha la facoltà di decidere liberamente come distribuire una certa somma tra sé e un altro partecipante. Tuttavia, se il secondo partecipante non accetta la proposta, entrambi restano a mani vuote. Com'è noto, i test mostrano che i partecipanti tendono a rifiutare le offerte repute ingiuste, agendo in contrasto con le norme di razionalità *standard* che prevedono la massimizzazione del profitto individuale⁶⁵. Le offerte ingiuste generano, infatti, un conflitto tra tendenze contrastanti: quella emotiva, mediata dall'insula anteriore bilaterale, porterebbe a resistere all'offerta, mentre quella razionale, mediata dalla corteccia prefrontale dorsolaterale, porterebbe ad accettare l'offerta⁶⁶.

Inoltre, gli studi sulla valutazione dell'iniquità procedurale e distributiva nell'allocazione delle risorse mostrano una marcata dissociazione dell'attivazione tra i due tipi di giudizio. Le valutazioni concernenti procedure inique suscitano una maggiore attivazione nelle aree cerebrali coinvolte nella cognizione sociale (corteccia prefrontale ventrolaterale, solco temporale superiore), mentre le valutazioni riguardanti gli esiti iniqui del processo decisionale suscitano una maggiore attivazione nelle aree coinvolte nell'elaborazione emotiva (corteccia cingolata anteriore, insula anteriore, corteccia prefrontale dorsolaterale)⁶⁷. Analogamente, il senso soggettivo di giustizia retributiva dipende dalla combinazione di valutazioni di tipo affettivo (nell'amigdala, nella corteccia prefrontale mediale e nella corteccia cingolata posteriore) relative alla gravità del crimine (quanto dovrebbe essere punito l'autore del reato?) e di tipo razionale-categorico (nella corteccia prefrontale dorsolaterale) circa la responsabilità individuale (l'autore del reato è responsabile o no?)⁶⁸.

Il coinvolgimento di processi affettivo-emozionali nella valutazione di varie forme di giustizia distributiva e retributiva non deve essere visto

⁶⁵ M.M. Pillutla, J.K. Murnighan, *Unfairness, Anger, and Spite: Emotional Rejections of Ultimatum Offers*, in «Organizational Behavior and Human Decision Processes», vol. 68, no. 3, 1996, pp. 208-224.

⁶⁶ A.G. Sanfey, J.K. Rilling, J.A. Aronson, L.E. Nystrom, J.D. Cohen, *The Neural Basis of Economic Decision-Making in the Ultimatum Game*, in «Science», vol. 300, no. 5626, 2003, pp. 1755-1758.

⁶⁷ J.H. Dulebohn, D.E. Conlon, I. Sarinopoulos, R.B. Davison, G. McNamara, *The Biological Bases of Unfairness: Neuroimaging Evidence for the Distinctiveness of Procedural and Distributive Justice*, in «Organizational Behavior and Human Decision Processes», vol. 110, no. 2, 2009, pp. 140-151.

⁶⁸ J.W. Buckholz, C.I. Asplund, P.E. Dux, D.H. Zald, J.C. Gore, O.D. Jones, R. Marois, *The Neural Correlates of Third-Party Punishment*, in «Neuron», vol. 60, no. 5, 2008, pp. 930-940.

semplificisticamente, in chiave esclusivamente negativa, come il risultato di pregiudizi e distorsioni cognitive che limitano la razionalità umana. Al contrario, questi meccanismi psicologici sono fondamentali per la formazione del nostro senso soggettivo di giustizia. Se è ingenuo attribuire forme di eccezionalismo ai professionisti del diritto per quanto concerne la permeabilità ai più comuni *bias* cognitivi⁶⁹, ancora più ingenuo parrebbe assumere che, nel prendere decisioni relative alla giustizia distributiva o retributiva, essi non si avvalgano a loro volta di questi medesimi processi che determinano una risposta di stampo cognitivo-affettivo, e non meramente razionalista, alle situazioni in essere. Guardando alla tradizione realista, ciò è fondamentale anche per come la giustizia funziona (e forse dovrebbe funzionare) in pratica – cioè in modi che *rispondono* al contesto specifico e alla soggettività di chi è giudicato piuttosto che basarsi unicamente su inferenze astratte relative alla sua identità sociale⁷⁰.

Inoltre, nell'enfatizzare la neutralità e l'uniformità prestazionale dell'I.A., è opportuno non confondere diverse forme di imparzialità: gli algoritmi potrebbero sì risultare più imparziali dei giudici umani nei confronti di specifici gruppi sociali. Questo non significa, però, che le decisioni dell'I.A. siano necessariamente imparziali rispetto ai sottostanti modelli e valori etici, sociali e politici. Al contrario, tali decisioni sono a loro volta impregnate di modelli e valori, anche se in modi che, ancora una volta, potrebbero risultare scarsamente comprensibili per i singoli: esse riflettono un compromesso specifico tra interessi potenzialmente contrapposti e, in ultima analisi, tra diversi modelli sociali e concezioni dell'equità, ad esempio circa come debba essere organizzata la cooperazione sociale. Queste differenze – che dipendono, in ultima analisi, da un'altra forma di discrezionalità, quella dei programmatori e dei *policy makers* – sono incorporate, ad esempio, in come le norme giuridiche, espresse nel linguaggio proposizionale ordinario, sono tradotte nei codici degli algoritmi; nei rischi e nei fallimenti epistemici che il sistema è disposto a tollerare quando prende decisioni sull'allocazione delle risorse o sull'appropriatezza di un certo trattamento; nei giudizi su quali problemi sociali richiedano più urgentemente un intervento; o nel modo in cui i singoli fattori peseranno sulle decisioni (per esempio, quanto il *background* socioeconomico di un imputato deve essere soppesato dagli strumenti di valutazione del rischio)⁷¹.

⁶⁹ G. Edmond, K. A. Martire, *Just Cognition: Scientific Research on Bias and Some Implications for Legal Procedure and Decision-Making*, cit.

⁷⁰ R. Dworkin, *Taking Rights Seriously*, London, Duckworth, 1977; R. Jorgensen, *Algorithms and the Individual in Criminal Law*, cit.

⁷¹ J. Biddle, *On Predicting Recidivism: Epistemic Risk, Tradeoffs, and Values in Machine*

A questo punto, è possibile delineare meglio i contorni del difficile compromesso tra discrezionalità e uniformità: da un lato, a patto di emanciparsi dal problema del pregiudizio algoritmico, l'I.A. promette di sbarazzarsi di forme di discrezionalità che sconfinano in una *cattiva* variabilità. Dall'altro lato, sarebbe pericoloso rinunciare a quella flessibilità e discrezionalità *buone* che, negli esseri umani, dipendono sia dall'integrazione di risorse e processi cognitivi diversi (emotivi e razionali) sia dall'adattamento flessibile ai cambiamenti culturali e alle circostanze specifiche.

È, tuttavia, opportuno evitare di scaderne in dicotomie semplicistiche opponendo la flessibilità dell'intelligenza umana all'uniformità dell'I.A. *In primis*, la discrezionalità potrebbe portare a essere oggetto di un trattamento peggiore o alla violazione di altri principi cardine della giustizia, come la necessità di applicare la legge in modo uguale per tutti o di distribuire in modo equo i vari oneri sociali⁷². In secondo luogo, la premessa secondo cui l'intelligenza umana è più flessibile di quella artificiale potrebbe suscitare un certo scetticismo. Ad esempio, in virtù della sua capacità di incrociare un numero maggiore di variabili, l'I.A. potrebbe anche rivelarsi più sensibile degli esseri umani alle sfumature del contesto. È poi discutibile se la relativa tendenza alla flessibilità o all'uniformità, laddove effettivamente esista, sia contingente e non inevitabile: se sia, quindi, impossibile in linea di principio trattare la risoluzione di casi giuridici nei termini di conclusioni derivabili da un set di premesse. Questa linea argomentativa, basata su una concezione di giustizia personalizzata che implica che casi analoghi possano essere trattati in modo diverso, sembra precludere in partenza la possibilità di affidarsi a un algoritmo per dirimere questioni di giustizia⁷³.

Infine, laddove si cerchi un compromesso tra i vantaggi dell'I.A. e forme di buona discrezionalità, le soluzioni ipotizzabili non sono esenti da difficoltà. Si pensi alla possibilità di dotare direttamente l'I.A. della capacità di esercitare una certa discrezionalità. Per funzionare in modo adeguatamente simile alla discrezionalità umana, questa dovrebbe essere quantomeno in grado di riflettere il consenso sociale, morale e legale su una certa questione, nonché evolvere in relazione ai cambiamenti della società e adattarsi a circostanze e compiti imprevisi⁷⁴. Per quanto si siano

Learning, cit.; M. Hildebrandt, *Algorithmic Regulation and the Rule of Law*, cit.

⁷² R. Jorgensen, *Algorithms and the Individual in Criminal Law*, cit.

⁷³ R. Binns, *Human Judgement in Algorithmic Loops; Individual Justice and Automated Decision-Making*, cit.

⁷⁴ R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, cit.

fatti notevoli progressi in questo senso grazie a forme di *machine learning* capaci di combinare un numero sempre più significativo di variabili, questa soluzione pone, tuttavia, considerevoli problemi a livello tecnico⁷⁵.

Altrettante questioni restano poi aperte sul piano normativo, in quanto non è chiaro come e quando questa flessibilità possa (o debba) essere implementata, fino a che punto le forme di discrezionalità buone e cattive possano essere nettamente disgiunte e quali aspetti giustifichino decisioni differenti per casi simili. Ciò vale, a maggior ragione, se si tengono in considerazione le inevitabili preoccupazioni relative al potenziale esercizio opaco della discrezionalità che un'I.A./scatola nera può offrire. L'alternativa forse più praticabile allo stato attuale potrebbe essere quella di sistematizzare maggiormente la divisione del lavoro tra giudici umani e I.A. – tenendo conto del fatto che la più parte delle situazioni in cui si richiede a esseri umani e I.A. di lavorare in tandem sono proprio quei casi in cui né i primi né la seconda sono in grado di garantire una *performance* nettamente superiore, tale da rendere superfluo l'apporto di uno o dell'altro attore⁷⁶. Questa divisione del lavoro dovrebbe trovare applicazione o attraverso una collaborazione fra giudici tradizionali e I.A. in diverse fasi del processo giudiziario relativamente ai medesimi casi o limitando l'uso dell'I.A. a casi selezionati. In virtù di questa collaborazione, gli esseri umani sarebbero, quindi, chiamati in causa ogniqualvolta la capacità di esercitare una certa discrezionalità sia considerata particolarmente preziosa. Tuttavia, determinare le forme e le modalità appropriate di una tale divisione del lavoro decisionale e chi debba definire quali casi meritino di essere giudicati con maggior flessibilità non è semplice. Peraltro, le situazioni in cui forme di cattiva discrezionalità sono da evitare sono spesso le stesse nelle quali forme di discrezionalità positiva, ovvero in sintonia con il contesto, sono elogiabili⁷⁷.

5. Conclusioni

In questo articolo ho discusso brevemente tre sfide etiche legate all'implementazione di sistemi di I.A. in varie forme di processi decisionali

⁷⁵ Si veda, ad esempio, N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lampos, *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*, in «PeerJ Computer Science», vol. 2, 2016, pp. 1-19; R. Binns, *Human Judgement in Algorithmic Loops; Individual Justice and Automated Decision-Making*, cit.

⁷⁶ J. Zerilli, U. Bhatt, A. Weller, *How Transparency Modulates Trust in Artificial Intelligence*, cit.

⁷⁷ R.M. Re, A. Solow-Niederman, *Developing Artificially Intelligent Justice*, cit.

legati alla giustizia: il problema della scatola nera, il problema dei pregiudizi umani e algoritmici e la questione del rapporto fra uniformità e discrezionalità nell'amministrazione della giustizia distributiva e retributiva. In particolare, anche presupponendo che le prime due questioni possano essere superate, la progressiva automatizzazione del sistema giudiziario potrebbe portare alla neutralizzazione delle buone forme di flessibilità e discrezionalità che sono più tipiche dell'intelligenza umana. Flessibilità e discrezionalità hanno a che fare con il diritto dell'agente di essere trattato come un individuo. Questo va oltre il diritto di essere giudicato da un arbitro imparziale e riguarda piuttosto il dovere del giudice o dell'impiegato di essere sensibile alla soggettività dei singoli e di rispettare la separatezza delle persone⁷⁸. Soprattutto se si è già in una posizione di vulnerabilità, essere trattati come individui include il diritto – potenzialmente eroso anche dall'opacità del processo decisionale algoritmico – di prendere parte al processo decisionale. Pertanto, trovare il giusto equilibrio tra uniformità e discrezionalità rappresenta una delle principali sfide nelle applicazioni giudiziarie dell'interazione fra esseri umani e I.A. Ciò naturalmente non toglie che, rispetto ai giudici umani, l'I.A. presenta, o potrebbe presentare in futuro, vantaggi concreti in termini di efficienza, riduzione dei tempi e dei costi e uniformità delle prestazioni. Tuttavia, il suo contributo al processo decisionale in ambito legale-giudiziario deve essere valutato con attenzione, visti i suoi potenziali svantaggi etici e l'impatto sui diritti umani fondamentali, come il diritto di essere giudicati davanti a un tribunale indipendente e di accedere a un sistema di giustizia penale basato sui diritti umani, la presunzione di innocenza o il rispetto della *privacy*. Ciò è tanto più vero in quanto gli attuali sistemi di I.A. stanno progressivamente passando da strumenti ausiliari a decisori veri e propri, in grado di incidere più direttamente sulla vita delle persone, formulando raccomandazioni, previsioni e deliberazioni.

⁷⁸ R. Jorgensen, *Algorithms and the Individual in Criminal Law*, cit.; K. Lippert-Rasmussen, 'We are all Different': *Statistical Discrimination and the Right to be Treated as an Individual*, in «J Ethics», vol. 15, 2011, pp. 47-59.

Responsabilità da danno algoritmico

MARTA INFANTINO*

SOMMARIO: 1. Introduzione – 2. Prospettive e caveat – 3. Lo spettro degli incidenti algoritmici – 4. Connotati e problemi – 5. Illeciti algoritmici e processi extracontrattuali – 5.1. Giudici e giurie – 5.2. Profili probatori – 5.3. Danni di massa e azioni collettive – 5.4. Il costo delle liti – 6. Conclusioni.

1. *Introduzione*

Il presente contributo si centra sugli ‘illeciti algoritmici’, definibili, in prima approssimazione, come incidenti nei quali un algoritmo più o meno intelligente, più o meno autonomo, provoca conseguenze dannose di natura extracontrattuale¹. Scopo dell’indagine, nel limitato spazio a disposizione, è illustrare come tali illeciti presentino caratteri, e sollevino problemi, che chiamano a una risposta differenziata a seconda della tradizione giuridica di riferimento. Un simile obiettivo rasenta ovviamente la banalità; eppure, esso merita di essere perseguito tenendo conto di come molti degli attuali dibattiti sulla regolazione algoritmica si orientino verso la ricerca di un regime di responsabilità che possa dirsi ottimale in ogni tempo e luogo, finendo così per sottovalutare l’importanza del contesto nel definire spazi e modi di intervento delle regole giuridiche.

A questo fine, dopo un breve inquadramento della prospettiva geografica e culturale nella quale il presente scritto intende collocarsi (par. 2), si offrirà qualche esemplificazione di cosa si intende qui per illecito algoritmico (par. 3), per poi analizzare gli elementi peculiari a tali illeciti e le sfide particolari che essi sollevano (par. 4). Si passerà quindi, sia pure molto in breve, a evidenziare come macro-regioni diverse del Nord del mondo siano

* DISPES, Università di Trieste.

¹ Questo breve saggio, che mantiene il tono colloquiale e la povertà di riferimenti propri all’originario intervento orale, si appoggia largamente a ricerche preesistenti, in particolare quelle condotte dall’autrice assieme al Dott. Weiwei Wang: si v., oltre che M. Infantino, *La responsabilità per danni algoritmici: prospettive europeo-continentali*, in «Responsabilità Civile e Previdenza», 2019, pp. 1762-1784, M. Infantino e W. Wang, *Algorithmic Torts. A Prospective Comparative Overview*, in «Transnational Law & Contemporary Problems», 28, 2019, pp. 309-362.

variamente attrezzate a reagire al moltiplicarsi di questi incidenti (par. 5): il richiamo è, in particolare, all'Europa continentale, alla Cina e agli Stati Uniti². Per verificare in quale misura i locali sistemi di responsabilità civile siano pronti ad affrontare la *litigation* extracontrattuale riguardo ai sinistri algoritmici, ci si soffermerà (non tanto sulle regole sostanziali inerenti il possibile regime di responsabilità, i principi di ricostruzione dell'eziologia dei danni, le poste risarcibili, le difese disponibili – questioni che sono solite attrarre, e hanno già attratto, ampia attenzione³ –, quanto) su aspetti di tipo procedurale, selezionati per la loro capacità di incidere sul raggio di operatività e l'amministrazione dei rimedi extracontrattuali. L'attenzione andrà allora a: la presenza, o no, di giurie entro i processi di responsabilità civile (par. 5.1), le regole probatorie (par. 5.2), la disponibilità di forme di aggregazione delle azioni individuali (par. 5.3) e la ripartizione dei costi di giustizia (par. 5.4). Ancorché una simile verifica verta su questioni assai limitate e frammentarie, e si mantenga a un livello di genericità tale da non rendere onore alle molteplici articolazioni interne esistenti entro le regioni di riferimento, i dati raccolti consentiranno di trarre alcune conclusioni sia riguardo alla variegata attitudine delle tradizioni in esame a reagire agli illeciti algoritmici, sia, più in generale, circa l'utilità, se non la necessità, di un approccio comparato al tema (par. 6).

² Quanto all'Europa, restano perciò esclusi dal raggio di analisi i sistemi di *common law* d'oltre Manica, e in particolare quello inglese. La scelta si deve alla circostanza che la commistione fra la famiglia giuridica di appartenenza – il *common law*, appunto – e la sedimentazione di regole provenienti dall'Unione Europea e dal sistema gravitante attorno alla Convenzione di Diritti dell'Uomo e delle Libertà Fondamentali rendono tali sistemi da un lato vicini all'esperienza statunitense, ma dall'altro lato anche partecipi di molti caratteri propri al continente. Onde evitare di complicare oltre modo il discorso, le pagine che seguono guarderanno dunque ai soli sistemi europeo-continentali, presi nel loro insieme. Per le medesime ragioni, si farà riferimento generico agli Stati Uniti, ancorché questi siano in effetti una federazione di cinquanta stati – più il sistema federale –, ciascuno dotato di un proprio diritto della responsabilità civile, e si prenderà in considerazione la sola Cina continentale, senza guardare alle giurisdizioni miste delle regioni amministrative speciali nelle isole di Macao e Hong Kong (il cui diritto privato, e quindi anche extracontrattuale, è rispettivamente di matrice sino-portoghese e sino-inglese).

³ In una letteratura già amplissima, si v., a titolo di esempio, E. Karner, B. A. Koch, M. Geistfeld, *Comparative Law Study on Civil Liability for Artificial Intelligence*, Publications Office of the European Union, 2021, <https://op.europa.eu/it/publication-detail/-/publication/8a32ccc3-0f83-11ec-9151-01aa75ed71a1/language-en/format-PDF/source-search>; nonché M. Infantino e W. Wang, *Algorithmic Torts*, cit., pp. 346-361.

2. Prospettive e caveat

Le società in cui viviamo sono oramai intrise di sistemi meno e più intelligenti che tengono traccia delle attività umane, gestiscono quantitativi enormi di informazioni, individuano correlazioni ed erogano servizi in modo rapido, seriale, costante e ininterrotto. I più intelligenti fra questi sistemi sanno riconoscere immagini, decrittare suoni, comporre sinfonie, dialogare con le persone e apprendere dal mondo in autonomia. È la c.d. intelligenza artificiale (d'ora in avanti IA), i cui impieghi correnti spaziano dal supporto delle diagnosi mediche all'internet-delle-cose (*Internet-of-things* o IoT), dalla robotica alla finanza, dalla sorveglianza alle vetture driverless.

Una simile rivoluzione è certamente suscettibile di produrre enormi benefici individuali e collettivi, ma anche di perturbare in modo dirompente gli assetti economici, politici e sociali del mondo analogico. In un contesto artificialmente intelligente, le applicazioni non solo contribuiscono a decidere quali malattie curare, in che settori investire e chi far arrestare, ma influiscono anche, con i loro mix di descrizione e predizione, sulle scelte riguardo i prodotti da acquistare, le cause da sposare, le persone da assumere, assicurare, ascoltare, votare.

Per quanto ci interessa, è chiaro che l'impiego massivo di algoritmi – se evita o comunque riduce il realizzarsi di molti incidenti tipicamente collegati alle attività umane – non elimina del tutto la possibilità che si verifichino contatti dannosi, e anzi apre la porta al possibile occorrere di tipologie inusate di sinistri. È perciò ragionevole supporre che, al moltiplicarsi delle attività algoritmiche, si accompagni il sorgere di nuovi incidenti⁴.

Si badi, tuttavia: quanto appena detto non significa affatto che la responsabilità extracontrattuale rappresenti la risposta ideale, né tanto meno l'unica, per regolare le odierne società algoritmiche. Il diritto del torto per definizione offre rimedi casuistici, bilaterali e successivi all'occorrere del danno, la cui attivazione è rimessa all'iniziativa individuale delle singole vittime (qualora queste ultime vogliano e possano farlo) e la cui concessione dipende dalle pronunce emesse caso per caso da corti civili variamente coordinati fra loro⁵. È poi dubbio che tali corti siano attrezzate

⁴ Si v., per tutti, W. Barfield e U. Pagallo, *Advanced Introduction to Law and Artificial Intelligence*, Cheltenham-Northampton, EE, 2020, pp. 93-95; D. Gifford, *Technological Triggers to Tort Revolutions: Steam Locomotives, Autonomous Vehicles, and Accident Compensation*, in «Journal of Tort Law», 11, 2018, pp. 71-143.

⁵ Non è certo questa la sede per approfondire i caratteri del diritto che ci occupa. Per una riflessione su funzioni e limiti del diritto della responsabilità civile in chiave storica e com-

a trattare complicate questioni tecnico-scientifiche e a assumere decisioni di rilevanza collettiva circa la distribuzione dei rischi tecnologici, specie considerato come – si sottolinea sovente⁶ – regole di responsabilità troppo severe potrebbero bensì proteggere i danneggiati, ma anche realizzare effetti contro-producenti per l'innovazione e l'economia nel suo complesso.

Resta però il fatto che, con tutte le loro imperfezioni, i sistemi di responsabilità civile si sono evoluti ed espansi, nel corso degli ultimi due secoli, di pari passo con l'industrializzazione e l'evoluzione tecnologica, fungendo giuridicamente ora da pungolo, ora da scudo, alla produzione di beni e allo svolgimento di attività intrinsecamente dannose⁷. Ecco perché, come si diceva, è ragionevole supporre che le regole di responsabilità civile avranno una parte importante da svolgere anche a fronte dell'attuale rivoluzione algoritmica. Non a caso, in tutte le regioni in esame, vi è un'esile ma crescente *litigation* extracontrattuale centrata sull'impiego di algoritmi⁸ e, specie in Europa continentale e negli Stati Uniti, si è già formato un dibattito piuttosto cospicuo circa il regime di responsabilità ottimale da applicare ai danni cagionati da algoritmi (soprattutto se integrati in un supporto fisico, come robot chirurgici e macchine a guida autonoma)⁹.

In questa sede, però, non ci interesseremo a quel dibattito, che è sovente teso a individuare il punto ottimale di compromesso fra copertura delle esternalità negative causate dall'attuale rivoluzione algoritmica e protezione del libero corso dell'innovazione tecnologica. Anzi, presupposto della presente analisi è proprio che non vi sia regola 'migliore' che possa dirsi tale in ogni luogo. Come si cercherà di dimostrare, ciascuna tradizione presenta dei propri macro-connotati giuridici, interni ed esterni al diritto

parata, si rimanda, per tutti, a M. Bussani, *L'illecito civile*, Napoli, ESI, 2020, pp. 113-160.

⁶ Si v., a questo riguardo, Commissione Europea, *Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale e modifica alcuni atti legislativi dell'Unione*, 2021/0106(COD), 21 aprile 2021, <https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>, considerando (3)-(5).

⁷ Si v. M. Bussani, *L'illecito civile*, cit., pp. 113-160; D. Gifford, *Technological Triggers to Tort Revolutions*, cit.

⁸ Si v. i casi citati *infra*, alle note 16, 18-21, 23.

⁹ Il dibattito è già troppo ampio per poter essere qui sintetizzato. Si rimanda alla letteratura citata in M. Infantino e W. Wang, *Algorithmic Torts*, cit., oltre che a GRECA (ed.), *La responsabilité civile et l'intelligence artificielle*, Brussels, Larcier, in corso di pubblicazione 2023; E. Karner, B. A. Koch, M. A. Geistfeld, C. Wendehorst (a cura di), *Civil Liability for Artificial Intelligence and Software*, Berlino, Boston, de Gruyter, 2022; A. Beckers e G. Teubner, *Three Liability Regimes for Artificial Intelligence. Algorithmic Actants, Hybrids, Crowds*, Oxford-Portland, Hart, 2021.

della responsabilità civile¹⁰, che incidono sullo spazio e sul raggio di operatività a questo riconosciuto e riconoscibile. In altri termini, non c'è una regola ottimale in astratto; ci possono essere solo regole più o meno adatte ai contesti nei quali, e alle finalità per le quali, queste sono chiamate a operare.

Un ultimo caveat. I macro-connotati che ci interessano, essendo radicati in profondità nell'architettura giuridica propria a ciascuna regione analizzata, sono tali da non poter essere modificati agilmente per effetto della semplice adozione, dall'alto, di regole nuove (per quanto 'migliori'). Sul punto, tutti i sistemi in esame sono all'erta. L'Ufficio Centrale per la Sicurezza Cibernetica della Repubblica Popolare Cinese ha adottato il 31 dicembre 2021 un "Regolamento sulla gestione dei sistemi algoritmici di raccomandazione", entrato in vigore in vigore il 1° marzo 2022, che impone regole di sicurezza e trasparenza agli operatori che impieghino sistemi personalizzati di suggerimento di informazioni, offerte e servizi¹¹. In Europa continentale, alla proposta di 'Regolamento sulla responsabilità per il funzionamento dei sistemi di intelligenza artificiale' presentata dal Parlamento europeo nell'ottobre 2020¹², si è affiancata quella elaborata dalla Commissione nell'aprile 2021 per un Regolamento c.d. Artificial Intelligence Act¹³. Negli Stati Uniti pende dal febbraio 2022 avanti al

¹⁰ Si pensi, per non fare che qualche esempio, alle regole che governano la circolazione dei dati, al ruolo ricoperto, in ciascun sistema, da amministrazioni, autorità indipendenti e assicurazioni private, al vario rapporto fra rimedi pubblicistici e privatistici, ai differenti tassi di accettazione dei rischi e di litigiosità propri alle varie culture giuridiche. Circa l'interdipendenza fra regole extracontrattuali e contesto giuridico, v., per tutti, M. Bussani e M. Infantino, *Tort Law and Legal Cultures*, in «American Journal of Comparative Law», 63, 2015, pp. 77-108.

¹¹ Ufficio Centrale per la Sicurezza Cibernetica, *Regolamento sulla gestione dei sistemi algoritmici di raccomandazione*, 31 dicembre 2021, a http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm.

¹² Parlamento Europeo, *Risoluzione del Parlamento europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale* (2020/2014(INL)), ottobre 2020, https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_IT.html.

¹³ Si v. retro, nota 6. Questa proposta, per la verità, non contiene regole di responsabilità ma istituisce un complesso di obblighi e forme di monitoraggio ex ante circa gli impieghi dell'intelligenza artificiale. Per un primo commento alla proposta, M. Ebers, V. R. S. Hoch, F. Rosenkranz, H. Ruschemeier, B. Steinrötter, *The European Commission's Proposal for an Artificial Intelligence Act—A Critical Assessment by Members of the Robotics and AI Law Society (RAILS)*, in «J», 4, 2021, pp. 589-603; L. Floridi, *The European Legislation on AI: A Brief Analysis of its Philosophical Approach*, in «Philosophy & Technology», 2021, pp. 215-222. Nel settembre 2022 la Commissione Europea ha pubblicato due ulteriori proposte (di direttiva), l'una centrata sull'ammodernamento delle norme in materia di responsabilità da prodotto

Congresso un disegno di legge intitolato ‘Algorithmic Accountability Act’¹⁴. Ma, occorre chiedersi, queste riforme sono o sarebbero in grado di modificare i fattori che, in ciascun ordinamento, determinano la quantità, la tipologia e i tassi di successo dei casi portati in giudizio? Il presupposto da cui parte il presente scritto è che la risposta a questa domanda sia, per lo meno nel breve periodo, largamente negativa. I tratti generali di un sistema – quelli che definiscono l’attitudine delle vittime di reagire ai torti, la facilità o la convenienza economica del farlo, la disponibilità dei giudici ad allargare le maglie della responsabilità – non sono elementi che si modificano rapidamente, né per effetto di un tratto di penna soltanto.

Ecco perché le seguenti pagine si centreranno, sia pure in modo necessariamente superficiale, a osservare alcuni aspetti apparentemente collaterali rispetto al diritto extracontrattuale, ma in realtà importanti e resistenti nel determinare la reattività dell’infrastruttura giuridica rispetto ai cambiamenti tecnico-sociali.

3. *Lo spettro degli incidenti algoritmici*

Gli illeciti algoritmici possono assumere una varietà infinita di forme.

Sono oramai note le conseguenze economiche massive che possono discendere dall’interconnessione fra molteplici robot finanziari. È rimasto celebre il c.d. Flash Crash del 6 maggio 2010, nel quale degli ordini errati di acquisto di azioni, ripetuti immediatamente da una moltitudine di *robo-advisors* ultraveloci e connessi, hanno generato un effetto a cascata tale da far crollare improvvisamente il mercato azionario statunitense e da bruciare centinaia di milioni di dollari in una ventina appena di minuti¹⁵.

Nell’ambito degli incidenti che determinano limitazioni e lesioni al

e l’altra mirante a introdurre alcuni correttivi probatori da applicare ai giudizi extracontrattuali di risarcimento dei danni cagionati da un’intelligenza artificiale: v. rispettivamente COM(2022)495 e COM(2022) 496.

¹⁴ H.R.6580 – Algorithmic Accountability Act of 2022, a <https://www.congress.gov/bill/117th-congress/house-bill/6580/text?r=2&s=1>. Come la proposta di regolamento avanzata dalla Commissione Europea (si v. retro, note 6 e 13), anche questo testo si limita a stabilire obblighi di comportamento e certificazione a carico delle imprese che ricorrono a forme di intelligenza artificiale.

¹⁵ U.S. Commodity Futures Trading Commission e U.S. Securities & Exchange Commission, *Findings Regarding the Market Events of May 6, 2010. Report of the Staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues*, 30 settembre 30, a <https://www.sec.gov/files/marketevents-report.pdf>.

corpo, sono altrettanti risaputi i pericoli associati agli algoritmi di polizia predittiva e a quelli per le diagnosi mediche: si immagini ad esempio un sistema intelligente che sia istruito a privilegiare la ricerca di criminali potenziali, anche laddove questo possa condurre al fermo di un innocente¹⁶, oppure a un programma costruito per supportare i medici nella valutazione della salute dei pazienti che, avendo tratto irragionevoli correlazioni fra certi sintomi e certe malattie, suggerisca esami e trattamenti non necessari, o viceversa non segnali esami e trattamenti necessari¹⁷.

Ma ancora, si pensi al costo economico e reputazionale che persone e imprese possono subire per una minima flessione del ranking dei risultati loro riferiti per effetto di minime modifiche negli algoritmi che muovono i motori di ricerca¹⁸, oppure in ragione delle funzioni di auto-completamento adoperate dagli stessi motori, che sovente associano soggetti a informazioni scorrelate, inventiere o spiacevoli¹⁹.

¹⁶ Negli Stati Uniti vi è oramai una copiosa giurisprudenza riguardo gli algoritmi di riconoscimento facciale, per lo più declinata in termini di azioni per ingiustificato arricchimento proposte da coloro le cui fotografie sono state inserite, senza autorizzazione, fra i dati volti a educare i sistemi intelligenti: *Vance et al v. Microsoft Corporation*, 14 aprile 2021, 534 F. Supp. 3d 1301 (US District Court for the Western District of Washington) e *Vance et al v. Amazon.com, Inc.*, 14 aprile 2021, 534 F. Supp. 3d 1314 (US District Court for the Western District of Washington). Similmente, in Cina (ma con riguardo a un software di creazione di personaggi elettronici, che aveva impiegato l'immagine dell'attore per creare un avatar a suo modello e poi – senza chiedere l'autorizzazione dell'interessato – concedere l'uso di quell'avatar a terzi), si v. Corte Suprema della Repubblica Popolare Cinese, *Tipiche cause civili di tutela giurisdizionale dei diritti della personalità dopo la promulgazione del codice civile*, 11 aprile 2022, no. 4 (la quale riporta una decisione di primo grado della Corte dell'internet di Pechino, a <https://www.court.gov.cn/zixun-xiangqing-354261.html>; occorre al riguardo notare che la Corte Suprema Cinese pubblica regolarmente liste di 'casi tipici', specialmente selezionati, al fine di offrirli quale guida ai tribunali).

¹⁷ A dispetto dei continui sviluppi dell'IA, i sistemi diagnostici intelligenti sono ancora ben lungi dall'essere accurati: si v. M. Mirbabaie, S. Stieglitz, N. R. J. Frick, *Artificial intelligence in disease diagnostics: A critical review and classification on the current state of research guiding future direction*, in «Health and Technology», 11, 2021, pp. 693-731.

¹⁸ Si v. il caso americano *Kinderstart.com v. Google, Inc.*, 16 marzo 2007, 2007 WL 831806 (N.D.Cal.) (ove, tuttavia, la corte ha accolto la difesa del convenuto Google, secondo il quale la scelta di ordine dei risultati di ricerca è espressione del diritto alla libertà di espressione, come tale coperto dal Primo Emendamento alla Costituzione americana).

¹⁹ Cfr. Prima Corte Intermedia di Pechino, *Ren Jiayu v. Beijing Baidu Technology & Science Co.*, 12 marzo 2015, 09558 (secondo la quale Baidu non risponde delle associazioni algoritmiche proposte dal suo motore di ricerca, poiché queste sono il frutto di un trattamento automatizzato di dati non imputabile all'impresa); Cour de Cassation, 1a sezione civile, 19 giugno 2013, n° 12-17.591, ECLI:FR:CCASS:2013:C100625 (che ha ritenuto Google irresponsabile delle associazioni diffamatorie proposte dalle sue funzioni di auto-comple-

Rilevano poi gli effetti discriminatori che possono colpire le persone quando postulano per un lavoro, un mutuo, un'assicurazione, grazie alle conclusioni tratte in virtù del trattamento statistico massivo delle informazioni loro e altrui²⁰, così come i portati negativi evidenti suscettibili di scaturire dall'automazione algoritmica dell'esercizio di funzioni pubbliche, come la ricerca di frodi ed evasioni fiscali e lo svolgimento di compiti amministrativi²¹.

In una società algoritmicamente avanzata, si deve anche tener conto delle conseguenze pregiudizievoli ricollegabili alla vulnerabilità dei sistemi

tamento in virtù dell'automatismo delle stesse); BGH, 14 maggio 2013, VI ZR 269/12, BGHZ 197, 213 (che invece ha ritenuto Google responsabile nel caso in cui il motore di ricerca proponga associazioni infondate e infamanti).

²⁰ Tali effetti sono ampiamente osservati e studiati con riguardo agli Stati Uniti: cfr. S. Zuboff, *The Age of Surveillance Capitalism*, New York, Public Affairs, 2019, pp. 3-4, 200-206; V. Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York, St. Martin's Press, 2018, pp. 1-12; C. O'Neil, *Weapons of Math Destruction*, New York, Crown, 2016, pp. 161-178. Ma si v. anche la decisione della Corte Suprema americana *TransUnion LLC v. Ramirez*, 25 giugno 2021, 141 S. Ct. 2190 (2021) (che ha precisato come individuare il gruppo di vittime legittimate a presentare un'azione collettiva contro un'agenzia di rating del credito privata che, utilizzando un algoritmo mal costruito, aveva erroneamente schedato circa 8.000 persone come potenziali terroristi).

²¹ Copiosi sono i casi al riguardo. Cfr., negli Stati Uniti, *K.W. v. Armstrong*, 28 marzo 2016, 180 F. Supp. 3d 703 (Idaho District Court) (sull'illegittimità dell'uso da parte dell'Idaho Department of Health & Welfare di un programma di budgeting che aveva erroneamente ridotto in modo automatizzato le sovvenzioni spettanti ai disabili di quello stato); *State v. Loomis*, 13 luglio 2016, 2016 WI 68 (Wisconsin Supreme Court) (circa la legittimità dell'impiego da parte dei tribunali penali di un algoritmo per la valutazione circa il pericolo di recidiva dei condannati). In Olanda, numerose decisioni hanno condannato vari enti amministrativi per l'impiego di algoritmi non trasparenti: Corte distrettuale dell'Aja, 2 marzo 2020, ECLI:NL:RBDHA:2020:865 (a essere contestato era un programma creato dall'amministrazione centrale per profilare le persone al fine di prevenire e combattere le frodi fiscali); Hoge Raad, 17 agosto 2018, 17/01448, ECLI:NL:HR:2018:1316 (sui programmi di calcolo del valore degli immobili ai fini della loro tassazione); Raad van State, 17 maggio 2017, ECLI:NL:RVS:2017:1259 (riguardo l'impiego di un algoritmo per automatizzare la concessione di autorizzazioni amministrative in materia ambientale). Simile giurisprudenza si è avuta in Italia, ove le corti amministrative hanno ripetutamente sottolineato l'illegittimità dell'impiego di algoritmi non trasparenti: Consiglio di Stato, 13 dicembre 2019, n. 8472, Foro it. 2020, III, 340 e Consiglio di Stato, 8 aprile 2019, n. 2270, in Foro it., 2019, III, 606 (circa l'utilizzo di un algoritmo di assegnazione automatica dei posti di insegnante, a concorso avvenuto). Sul punto si è pronunciato anche il Consiglio costituzionale francese, ribadendo che l'amministrazione può fare ricorso ad algoritmi solo in alcuni settori, in modo trasparente e garantendo la possibilità di contraddittorio con gli interessati: Conseil constitutionnel, 12 giugno 2018, no. 2018-765, a <https://www.conseil-constitutionnel.fr/decision/2018/2018765DC.htm>.

informatici a black-out energetici, cyber-attacchi e altri impieghi malevoli²², nonché persino delle ipotesi in cui il danno deriva non già dal ricorso ad algoritmi, bensì dal loro mancato uso: si registrano già, negli Stati Uniti, azioni risarcitorie per *medical malpractice* avanzate contro medici e ospedali per l'omesso utilizzo di robot chirurgici, laddove la messa in campo di questi ultimi – a dire degli attori – avrebbe evitato l'errore umano di cui essi sono stati vittime²³.

La lista dei possibili incidenti, ovviamente, potrebbe continuare a lungo, ma il magro elenco appena svolto vale già a mettere in luce la quantità e l'impatto delle esternalità negative che possono essere prodotte dagli illeciti algoritmici. Si tratta ora di capire meglio perché la categoria merita rilievo autonomo e perché i problemi che essa solleva sono peculiari e distinti da quelli noti ai paesi industrializzati.

4. *Connotati e problemi*

Fra le fattispecie dannose sopra richiamate, vi sono in effetti degli elementi comuni che hanno un preciso impatto sull'attitudine dei meccanismi extracontrattuali tradizionali di far loro fronte²⁴.

Anzitutto, quale che sia il linguaggio di programmazione impiegato, qualsiasi algoritmo richiede competenze specializzate per essere (scritto e) letto. Quanto più complesso e autonomo è poi l'algoritmo, tanto più difficile è, anche per chi è esperto, comprendere e ricostruire il suo modo di operare (fino al livello massimo di incomprensibilità, proprio ai sistemi

²² Per l'analisi delle molteplici questioni sollevate dalla sicurezza dell'IA, si v., per tutti, R. Walters e M. Novak, *Cyber Security*, in Robert Walters e Marko Novak (a cura di), *Cyber Security, Artificial Intelligence, Data Protection & the Law*, Singapore, Springer Nature, 2021, pp. 21-37.

²³ Si v. i casi americani (entrambi di rigetto delle domande per difetto di prova di un nesso causale adeguato) *O'Brien v. Intuitive Surgical, Inc.*, 15 luglio 2011, WL 3040479 (Northern District Illinois); *Mracek v. Bryn Mawr Hosp.*, 28 febbraio 2010, 363 Fed. App'x 925 (3rd Circuit). Questi casi sono interessanti, perché rovesciano il tradizionale metro di paragone nelle azioni di responsabilità: invece che chiedersi se l'algoritmo si sia comportato in modo ragionevole, considerato ciò che avrebbe fatto al suo posto un essere umano, ci si chiede se il convenuto (umano) si sia comportato ragionevolmente, tenuto conto di cosa avrebbe fatto al suo posto l'algoritmo.

²⁴ Su questi caratteri, si v. variamente W. Barfield e U. Pagallo, *Advanced Introduction to Law and Artificial Intelligence*, cit., pp. 1-16; M. Infantino e W. Wang, *Algorithmic Torts*, cit., pp. 317-320.

c.d. black-box).

Due, molti algoritmi hanno origini molteplici e intricate, nel senso che sono il risultato della sedimentazione del lavoro di molte persone – società big tech, piccole start-up, sviluppatori in-house, programmatori indipendenti – le quali contribuiscono a un frammento del risultato finale, sovente in assenza di qualsiasi coordinazione fra loro.

Tre, la qualità di un algoritmo dipende ovviamente sia dalla qualità delle istruzioni che esso riceve, sia dalla qualità dei dati che gli sono forniti (o sui quali viene formato). Ma le istruzioni per definizione non sono neutrali, perché sono necessariamente influenzate dalle categorie, dalle visioni e dai pregiudizi di chi le formula, e i dati sono raramente completi, precisi, aggiornati, rappresentativi, perché anch'essi sono frutto di selezioni operate dai programmatori o, assai più spesso, da terzi (inclusi altri algoritmi).

Quattro, gli algoritmi non vivono in isolamento, ma solitamente sono interconnessi ad altri algoritmi (la macchina autonoma parla con lo smartphone del guidatore, con gli altri veicoli, la strada, i satelliti, e così via); spesso sono anche integrati in supporti fisici, nel qual caso la loro affidabilità dipende anche dalle condizioni materiali e dal tasso di funzionamento di quei supporti.

Cinque, le decisioni algoritmiche sono di regola meccaniche e generalizzate, nel senso che, nel momento in cui vengono assunte, incidono istantaneamente e immediatamente sulla situazione delle persone cui si applicano. Data la propensione degli algoritmi di circolare attraverso le frontiere, questa capacità di incidere simultaneamente e direttamente sugli interessati riguarda sovente soggetti siti in molti luoghi diversi fra loro, e spesso anche assai lontani da quelli ove si trovano gli originari programmatori/produttori/utilizzatori.

Ognuno di questi aspetti crea problemi specifici alla capacità del diritto della responsabilità civile di operare²⁵.

La complessità tecnica e l'opacità dei codici interni agli algoritmi, la loro dipendenza da istruzioni, dati, altri algoritmi e prodotti fisici, le loro possibili capacità di auto-apprendimento, sono tutti elementi che rendono assai difficile e costosa, per le vittime, l'identificazione della fonte dell'incidente e del soggetto che possa essere chiamato a risponderne (anche ammesso, il che non deve essere dato per scontato, che i danneggiati riescano a rendersi conto che qualcosa è andato storto). Il problema è nel

²⁵ Per una più approfondita analisi delle difficoltà sollevate dagli illeciti algoritmici nell'ottica extracontrattuale, cfr. E. K. Bernhard A. Koch, M. Geistfeld, *Comparative Law Study on Civil Liability for Artificial Intelligence*, cit.; M. Infantino e W. Wang, *Algorithmic Torts*, cit., 320, pp. 336-361.

codice? Nelle istruzioni o nei dati offerti all'algoritmo? Nell'interazione di questo con altri algoritmi o prodotti? Nel modo in cui è stato impiegato da un essere umano? La ricerca dell'eziologia precisa dell'incidente dovrà di necessità farsi con l'ausilio di esperti, capaci di leggere i codici, e con avvocati specializzati, che siano in grado di comprendere e tradurre quell'expertise tecnica in modo comprensibile alle corti (e alle giurie, là dove presenti). Esperti e avvocati vanno pagati, il che di conseguenza richiede la disponibilità di denaro sufficiente a finanziare la lite.

Non solo. In ragione della sostanziale ubiquità e dell'uso massivo di algoritmi, un incidente algoritmico è potenzialmente suscettibile di causare un ampio spettro di pregiudizi (fisici, non patrimoniali, economici), inquadrabili in una varietà di azioni proponibili nei confronti di una molteplicità di figure – dalla responsabilità per colpa nei riguardi di chi ha scritto o impiegato l'algoritmo, alla responsabilità da prodotto contro chi l'ha concepito o commercializzato, alla responsabilità pubblica o per lesione di diritti fondamentali nei confronti delle autorità che hanno mancato di verificarne l'operare, e così via. Scegliere contro chi agire e in base a quale titolo impone determinazioni complesse e rischiose, che devono necessariamente essere effettuate da professionisti specializzati. Complessità e rischiosità si moltiplicano quando, in ragione dell'attitudine degli algoritmi a viaggiare oltre le frontiere, l'incidente possa qualificarsi come illecito transnazionale. In mancanza, per il momento, di una regolazione globale sul punto, tale qualifica apre una serie di problemi ulteriori riguardo all'individuazione della giurisdizione più appropriata per radicare la lite e del diritto in base al quale quest'ultima sarà decisa (così ovviamente anche stimolando strategie di concorrenza regolatoria e di forum shopping). Difficoltà ulteriori sorgono poi allorché si tratti di sinistri che coinvolgono schiere cospicue di vittime, ove cruciale diventa la disponibilità, eventualmente anche transfrontaliera, di istituti per l'aggregazione o il coordinamento delle azioni individuali.

Insomma, gli illeciti algoritmici sollevano un gran numero di sfide ai potenziali attori, per i quali può diventare molto arduo dimostrare cosa ha recato loro danno e come, determinare quale possa essere il convenuto e la giurisdizione appropriata per instaurare un litigio, raccogliere le prove necessarie a sostenere la domanda, riunire assieme fasci di azioni, nonché ottenere forme di finanziamento alle liti in grado di sostenere tutti questi passaggi.

5. *Illeciti algoritmici e processi extracontrattuali*

Di qui la domanda al cuore di questo scritto: in quale misura le diverse tradizioni giuridiche qui in esame sono attrezzata a reagire a simili sfide?

Contrariamente a quanto usualmente assunto dai dibattiti in materia, la risposta a tale quesito (non solo varia da luogo a luogo, ma anche) non dipende esclusivamente dal regime di responsabilità, attuale o prospettico, che si voglia applicare agli incidenti algoritmici. Viceversa, essa è una variabile dipendente da una molteplicità di fattori, inclusi gli istituti e i principi procedurali che governano le controversie extracontrattuali.

È proprio su alcuni di questi che si concentrerà la nostra attenzione. Valuteremo in particolare come la compresenza di giudici togati e giurie (par. 5.1), le regole riguardo l'accesso e la valutazione delle prove scientifiche (par. 5.2), la disponibilità, più o meno ampia, di azioni collettive (par. 5.3) e i sistemi di ripartizione dei costi processuali (par. 5.4), siano tutti elementi apparentemente collaterali rispetto al diritto extracontrattuale, epperò suscettibili di incidere sulla capacità delle vittime di reagire agli incidenti algoritmici.

Ovviamente, si tratta solo di una selezione fra i tanti caratteri passibili di rilevare nella nostra materia; altrettanto ovviamente, ciascuno dei fattori appena ricordati è esposto a possibili mutamenti nel corso del tempo. È bene però ribadire, come già sottolineato retro, al par. 2, che da un lato la più parte delle proposte di riforma per ora avanzate si è centrata sulla sostanza, piuttosto che sulla procedura, dei giudizi di responsabilità, e che, dall'altro lato, cambiamenti macroscopici del modo di operare dei sistemi di responsabilità sono tanto più improbabili nel breve termine quanto più radicati e pervasivi sono i caratteri di cui si discute.

5.1. *Giudici e giurie*

La possibile presenza di una giuria laica a fianco del giudice è un connotato caratteristico del *tort law* americano.

Tale connotato ha conseguenze macroscopiche sull'amministrazione statunitense del diritto del torto, influenzando il tipo di argomenti impiegati avanti alle corti e lo stile di loro presentazione, favorendo l'oralità e la concentrazione della procedura, imponendo una notevole spettacolarizzazione nella gestione processuale e soprattutto incidendo sui tassi di successo e sul valore delle condanne. È difatti noto come le

giurie tendano a essere notevolmente generose nei confronti delle vittime e quindi contribuiscano sovente all'adozione, tipicamente in prima istanza, di decisioni ove le somme dovute dal convenuto a titolo di risarcimento compensativo ed eventualmente punitivo sono quantificate in importi assai elevati. Sebbene queste decisioni raramente sopravvivano nei gradi successivi (e nonostante l'indubbia diminuzione, specie durante il recente momento pandemico, dei tassi di ricorso alle giurie), questo tratto resta un elemento tipico, oltre che un enorme fattore di attrattività, della giurisdizione extracontrattuale americana²⁶.

Altrettanto noto è come le giurie siano semi-sconosciute alla procedura civile europeo-continentale e cinese. Ciò ovviamente riduce le aspettative delle vittime europee e cinesi di ottenere condanne risarcitorie dall'importo consistente, anche in virtù dell'impossibilità (in Europa) e dei ridotti limiti (in Cina) di concessione di danni punitivi, ma le assicura pure della circostanza che gli elementi tecnici del giudizio – sia tecnologici che giuridici – siano sempre valutati da giudici professionisti²⁷.

In altri termini, il monopolio del giudice minimizza il rischio, solitamente elevato con le giurie, di incomprensioni del diritto e delle risultanze tecniche, rafforzando in questo modo la predicibilità complessiva del sistema²⁸.

5.2. *Profili probatori*

Chiunque intenda far valere in giudizio una pretesa extracontrattuale deve ovviamente poter provare non solo di aver subito un danno ma anche di averlo subito a causa di un'azione, omissione o comunque evento ricollegabile al convenuto. I meccanismi probatori sono perciò cruciali rispetto alla proponibilità delle azioni di torto.

Al riguardo è anzitutto fondamentale ricordare l'istituto statunitense della *pre-trial discovery*, ossia il meccanismo procedurale (antecedente

²⁶ Su tutto ciò, M. Bussani, A. J. Sebok, M. Infantino, *Common Law and Civil Law Perspectives on Tort Law*, Oxford, OUP, 2022, 8-9; U. Magnus, *Why is US Tort Law so Different?*, in «Journal of European Tort Law», 1, 2010, pp. 102 – 124.

²⁷ Se sul suolo europeo non vi è spazio per un'espressa pronuncia dai connotati punitivi, questa è consentita dalle regole cinesi in materia di responsabilità da prodotto e protezione dei consumatori (si v. ad esempio l'art. 1207 del Codice civile cinese). Le corti cinesi, tuttavia, sono state finora piuttosto restie a sfruttare appieno tale possibilità: A. Janssen e J. Wang, *Punitive damages under the new Chinese Civil Code – a critical and comparative analysis*, 29 in «Asia-Pacific Law Review», 29, 2021, pp. 346-365.

²⁸ M. Infantino e W. Wang, *Algorithmic Torts*, cit., pp. 341-342.

all'instaurazione del giudizio) che consente a una parte di chiedere alla propria controparte di consegnarle o comunque darle accesso a tutta la documentazione rilevante per il caso. Tale potente istituto non esiste né in Europa continentale né in Cina, dove gli attori in giudizio, se intendono acquisire documenti o prove in possesso dell'altra parte, devono persuadere il giudice a ordinarne l'acquisizione²⁹.

Il rovescio della medaglia, in Europa come in Cina, è però che, nel caso di processi che riguardino complesse prove tecniche – come normalmente è il caso per gli incidenti algoritmici –, gli attori possono fare affidamento (non solo sulle risultanze ottenute da autorità pubbliche nel corso delle loro investigazioni amministrative o penali, ma anche) sugli ampi poteri del giudice di nominare un consulente tecnico neutrale che investighi gli aspetti principali della lite. Il punto marca una distanza notevole rispetto agli Stati Uniti, dove gli esperti principali sono invece quelli di parte (con l'effetto di trasformare la controversia in una battaglia di esperti, pagati dalle parti) e il potere del giudice di nominarne uno neutrale di ufficio è residuale³⁰.

In altri termini, le consulenze tecniche nei giudizi statunitensi sono solitamente poco affidabili nel merito (perché faziose e parziali) ed estremamente costose. Anche su questo fronte, le regole procedurali europee e cinesi privilegiano maggiormente la certezza del diritto e l'accessibilità dei rimedi, nella misura in cui consentono di contenere il costo delle dispute.

5.3. *Danni di massa e azioni collettive*

Ulteriore peculiarità della procedura civile statunitense è la disponibilità di azioni collettive, ancorché l'effettiva praticabilità di tale rimedio, specie per quanto attiene le cause che coinvolgano persone e diritti stranieri, sia negli ultimi anni andata costantemente diminuendo³¹.

Aggregare fasci di azioni individuali in una singola lite è uno strumento

²⁹ Cfr. E. Fahey e Z. Tao, *The Pretrial Discovery Process in Civil Cases: A Comparison of Evidence Discovery Between China and the United States*, in «Boston College International & Comparative Law Review», 37, 2014, pp. 281-332; Ulrich Magnus, *Why is US Tort Law so Different?*, cit., pp. 116-117.

³⁰ P. Monaco, *Sostanze tossiche e danni. Profili di diritto globale, europeo e nazionale*, Napoli, Editoriale Scientifica, 2020, pp. 77-98; M. Infantino e W. Wang, *Algorithmic Torts*, cit., pp. 342-343; E. Fahey e Z. Tao, *The Pretrial Discovery Process in Civil Cases*, cit., pp. 289-291.

³¹ M. Bussani, A. J. Sebok, M. Infantino, *Common Law and Civil Law Perspectives on Tort Law*, cit., 19; P. Monaco, *Sostanze tossiche e danni*, cit., pp. 139-153; U. Magnus, *Why is US Tort Law so Different?*, cit., pp. 115-116.

che consente agli attori di minimizzare costi e rischi dell'agire in giudizio, e che risulta particolarmente prezioso nel caso di incidenti algoritmici. Come accennato retro, al par. 4, questi ultimi sovente producono danni di massa, ossia colpiscono schiere indeterminate di persone, ciascuna delle quali soffre però un nocumento di ammontare molto basso o addirittura nullo: il caso paradigmatico è quello di algoritmi che violano la privacy o che hanno l'effetto di discriminare nei confronti di specifiche categorie di persone, ove i riverberi negativi individuali possono essere evanescenti, anche a fronte di un pregiudizio collettivo o sociale molto elevato³².

Ora, le azioni collettive sono tradizionalmente estranee tanto all'Europa quanto alla Cina, sebbene in entrambe le regioni sia evidente la tendenza recente a trapiantarne ed estenderne l'uso³³.

Occorre però anche qui porre attenzione: se è indubbio che, in Europa come in Cina, le azioni collettive sul fronte privatistico sono indisponibili o comunque restano poco utilizzate, è tuttavia altrettanto certo che, là dove sono in gioco danni di massa, i danneggiati europei e cinesi possono contare sull'attivismo delle procure e l'interventismo di autorità indipendenti e amministrazioni pubbliche (oltre che dello stesso legislatore) assai più di quanto non possano fare i loro omologhi statunitensi³⁴.

5.4. *Il costo delle liti*

Cruciale, nella scelta delle vittime se tentare, o no, un'azione giudiziale, è anche la questione circa i costi della lite.

Negli Stati Uniti come in Cina, le parti di una disputa di responsabilità civile pagano ciascuna i propri avvocati, sia che questi vincano la causa, sia che la perdano³⁵.

³² Non è caso che molte delle azioni per incidenti algoritmici, negli Stati Uniti, prendano la forma di azioni collettive: si v. i casi *TransUnion LLC v. Ramirez*, 25 giugno 2021, cit.; *K.W. v. Armstrong*, 28 marzo 2016, cit.

³³ Si v., ad esempio, in Europa, la recente Direttiva (UE) 2020/1828 del Parlamento europeo e del Consiglio del 25 novembre 2020 relativa alle azioni rappresentative a tutela degli interessi collettivi dei consumatori, e, in Cina, gli articoli 52-55 della Legge sulla procedura civile del 2017; v. anche A. Stadler, E. Jeuland, V. Smith (a cura di), *Collective and Mass Litigation in Europe. Model Rules for Effective Dispute Resolution*, Cheltenham-Northampton, EE, 2020; R. H. Huang, *Class Action in China. Challenges and Opportunities*, in Brian T. Fitzpatrick (a cura di), *The Cambridge Handbook of Class Actions. An International Survey*, Cambridge, CUP, 2021, pp. 350-365.

³⁴ M. Infantino e W. Wang, *Algorithmic Torts*, cit., p. 344.

³⁵ M. Reimann, *General Report*, in M. Reimann (a cura di), *Cost and Fee Allocation in Civil Procedure*, London-New York, Springer, 2012, 3, pp. 9-16; J. Xiao e X. Tang, *Cost and Fee*

In linea di principio, ciò significa che gli attori sopportano un rischio economico considerevole. La regola è tuttavia contemperata, in particolar modo sul suolo statunitense, dalla possibilità di scaricare parte del rischio della lite sullo stesso avvocato attraverso la stipulazione di un'apposita *contingent fee*, in base alla quale l'avvocato riceve zero se perde ma ottiene una quota di quanto riconosciuto a titolo di risarcimento se vince – tipicamente, un terzo³⁶. Questo fa sì che le cause di valore e chances di successo elevate trovino negli Stati Uniti quasi sempre un avvocato disposto ad assumersi il rischio, via *contingent fees*, di portarle in giudizio.

Ben diverso è il quadro in Europa, dove l'allocazione dei costi processuali segue largamente l'opposta regola del '*loser pays*' (e dove storicamente le *contingent fees* sono vietate o poco utilizzate): chi perde paga anche i costi della difesa sopportati dall'altra parte³⁷.

Si potrebbe quindi essere tentati di concludere che le vittime europee si trovino in una condizione largamente peggiore rispetto ai loro simili negli USA. È però necessario porre attenzione ad almeno due dati ulteriori.

Il primo punto è che sia l'Europa che la Cina hanno ben sviluppati programmi pubblici di gratuito patrocinio, i quali restano sostanzialmente sconosciuti agli Stati Uniti, al di fuori della limitata area di intervento della c.d. *pro bono litigation*³⁸. In altri termini, chiunque, in Europa o in Cina, abbia un valido motivo di azione, pure in controversie dal valore relativamente basso, ha almeno in teoria la possibilità di accedere al sistema giudiziario; è viceversa assai improbabile che quella medesima persona possa trovare un avvocato interessato a difenderlo negli Stati Uniti³⁹.

Il secondo aspetto che merita il rilievo concerne l'ammontare dei costi processuali, latamente intesi – inclusivi cioè non solo delle spese per gli avvocati, ma anche delle tasse e di qualsiasi altra contribuzione necessaria a portare un litigio avanti una corte. Quando valutati rispetto al valore delle cause e al costo della vita, i costi cinesi risultano essere inferiori a quelli europei, e i dati di entrambe le regioni restano incomparabilmente più bassi rispetto ai loro equivalenti americani⁴⁰. Anche sotto tale profilo, pertanto,

Allocation in Civil Procedure: China National Report, in «Tsinghua China Law Review» 4, 45, 2012, pp. 46-49.

³⁶ M. Reimann, *General Report*, cit., p. 45.

³⁷ Ivi, cit., pp. 9-16.

³⁸ Ivi, cit., 3, pp. 36-38; J. Xiao e X. Tang, *Cost and Fee Allocation in Civil Procedure: China National Report*, in «Tsinghua China Law Review», 4 45, 2012, pp. 61-65; U. Magnus, *Why is US Tort Law so Different?*, cit., pp. 114-115.

³⁹ U. Magnus, *Why is US Tort Law so Different?*, cit., pp. 113-114.

⁴⁰ M. Henssler, *The Organization of Legal Professions*, in M. Schauer e B. Verschraegen (a

la giustizia civile in Cina e in Europa appare maggiormente accessibile di quanto non sia quella americana.

6. Conclusioni

Cosa può trarsi dal frammentario quadro finora tracciato?

Ogni tradizione ha regole, istituti e prassi, di natura sostanziale e procedurale, ben impressi nel DNA del proprio diritto della responsabilità civile. Non si sono qui menzionati aspetti sostanziali, ma si è guardato solo a (una selezione di) elementi procedurali.

Persino un'analisi limitata come quella fin qui svolta, però, vale a dimostrare la validità della tesi avanzata in esergo: anche questioni apparentemente lontane dal cuore del diritto extracontrattuale possono avere un'importanza fondamentale nel determinare lo sviluppo di quest'ultimo – in generale, e con riferimento agli illeciti algoritmici in particolare.

Si è ad esempio visto che gli Stati Uniti si offrono quale giurisdizione più accogliente per le vittime che abbiano danni ingenti, singolarmente o collettivamente, data la facilità, in questi casi, di aggregare assieme più azioni e di trovare avvocati disposti a finanziare le liti, anche grazie al miraggio delle elevate condanne (eventualmente anche punitive) che le giurie possono riconoscere in primo grado.

Si è tuttavia anche messo in luce come molti connotati procedurali tipici del diritto americano – la presenza delle giurie, la disponibilità di strumenti di pre-trial discovery, le battaglie fra esperti – in pratica rendono la *litigation made in USA* imprevedibile e costosa.

Per contro, sul continente euroasiatico, l'accesso alla giustizia è tendenzialmente più diretto e semplice, anche per chi ha sofferto danni non gravi o ingiustizie poco eclatanti. Ancorché vi siano notevoli difficoltà a incentivare la proposizione di litigi meritevoli, specie nel caso di incidenti di massa, entrambi i sistemi giudiziari risultano nel suo complesso più stabili e prevedibili del loro omologo oltre-atlantico.

Gli elementi appena evidenziati parrebbero suggerire che le vittime di incidenti algoritmici abbiano un interesse maggiore ad agire in Europa o in

cura di), *General Reports of the XIXth Congress of the International Academy of Comparative Law*, Cham, Springer, 2017, pp. 261, 277; C. Hodges, S. Vogenauer, M. Tulibacka, *The Oxford Study on Costs and Funding of Civil Litigation*, in C. Hodges, S. Vogenauer, M. Tulibacka (a cura di), *The Costs and Funding of Civil Litigation: A Comparative Perspective*, Oxford-Portland, Hart, 2010, 3, pp. 69-71.

Cina piuttosto che negli Stati Uniti, e che viceversa ai convenuti, specie se in forma *corporate*, dovrebbe in principio essere più congeniale la giurisdizione statunitense, ove essi potrebbero giovare degli alti costi del processo e della disponibilità di consulenti esperti al soldo del miglior offerente. Ma prevedere come si muoverà in futuro la *litigation* sul punto è compito che richiederebbe studi, ricerche e competenze di ben altro calibro.

Un solo punto è sicuro: i meccanismi extracontrattuali di ripartizione sociale del costo dei danni algoritmici seguono linee che non dipendono esclusivamente da, né si possono cambiare semplicemente con, macro-scelte effettuate una volta e per sempre dall'alto. Ugualmente importanti, nel settore che ci occupa come in molti altri, sono microelementi contestuali che si muovono dal basso e dal basso incidono, impercettibilmente ma potentemente, sul comportamento complessivo del sistema. A importare cioè è anche quanto spesso, troppo spesso, resta invisibile agli occhi (dei giuristi come dei riformatori).

Il diritto della modernità, dall'età delle codificazioni in avanti, si è sviluppato assumendo una certa idea di soggetto di diritto, di *homo juridicus*: un soggetto razionale, pienamente capace di intendere e di volere e dunque "responsabile" per i propri atti, leciti e illeciti – salva la dimostrazione dell'assenza, intesa come fatto eccezionale, come eccezione alla regola, di tale piena capacità. L'*homo juridicus* è un soggetto che sa ciò che vuole, e che è in grado di operare un attento calcolo costi-benefici tra un certo corso di azione e le conseguenze ad esso associate dall'ordinamento giuridico. E tali capacità razionali, in fondo, sono ciò che ci rende tutti uguali davanti alla legge – tutti, parimenti, soggetti di diritto. Questo paradigma – il paradigma del soggetto razionale – appare per molti versi in crisi, la sua dimensione illusoria diventa sempre più evidente. E questa crisi è esplorata, nella prospettiva di una riflessione multidisciplinare, nei contributi raccolti in questo volume, ponendo sotto esame questioni come: il rapporto tra sentimenti, emozioni e razionalità nei giudizi morali; il ricorso alle neuroscienze per esplorare la nozione di "libera determinazione" da parte del soggetto agente; la problematica configurabilità del concetto di libero arbitrio, e le conseguenze di ciò per l'imputabilità della responsabilità penale; il ruolo delle emozioni tanto nella realizzazione della condotta penalmente rilevante, quanto nella valutazione che di tale condotta è chiamato a fare il giudice; le questioni etiche e giuridiche associate alle decisioni prese da un'intelligenza artificiale.

Antonio Carratta

è Professore di Diritto processuale civile presso il Dipartimento di Giurisprudenza dell'Università Roma Tre.

Mario De Caro

è Professore di Filosofia morale presso il Dipartimento di Filosofia, Comunicazione e Spettacolo dell'Università Roma Tre.

Giorgio Pino

è Professore di Filosofia del diritto presso il Dipartimento di Giurisprudenza dell'Università Roma Tre.