

IL PROBLEMA DEI PARADOSSI MORALI E LA TEORIA DELL'ERRORE: DUE PERCORSI E UN INCROCIO

GIUSEPPE ROCCHÈ

*Dipartimento di Giurisprudenza
Università di Palermo
giuseppe.rocche@unipa.it*

ABSTRACT

The sphere of moral action is plagued by paradoxes. Bruno Celano considers some traditional moral paradoxes related to recursive moral norms and tragic choices. This article extends Celano's analysis to other paradoxes discussed in recent decades in the context of the debates on population ethics and the Trolley Problem. The prospect that there is no moral theory capable of respecting our moral intuitions can be seen as a threat to moral realism. This idea was recently developed by Christopher Cowie in his original defence of the moral error theory. A distinctive feature of Cowie's view is the connection between normative ethics and metaethics. This article proposes a second way of using the problem of moral paradoxes to connect normative enquiry and metaethics, the structure of which is opposite to Cowie's: not "from ethics to metaethics", but "from metaethics to ethics".

KEYWORDS

Christopher Cowie; Population ethics; Moral error theory; Bruno Celano; Moral paradoxes.

1. BRUNO CELANO E LA CRISI DEL DIRITTO NATURALE

Nella sua introduzione allo studio del diritto Bruno Celano colloca una riflessione sul rapporto tra ragionamento giuridico e ragionamento morale per certi versi sorprendente.

Frequentemente il rapporto tra diritto e morale vede il ragionamento giuridico attingere dal repertorio delle considerazioni morali. La morale che va in soccorso del diritto. Questo ruolo della morale è centrale, ad esempio, negli scritti di Ronald Dworkin sulla lettura morale del diritto e in particolare della costituzione americana¹. Celano osserva come il rapporto tra diritto e morale possa però presentarsi

¹ Dworkin 1996.

anche in modo invertito. Non la morale che va in soccorso del diritto, ma il diritto che va in soccorso della morale². A rendere intellegibile questo rapporto è una rappresentazione tetra del ragionamento morale.

Celano utilizza l'Orestea di Eschilo come scenario illustrativo della sua idea. Oreste torna a casa per vendicare la morte di suo padre Agamennone e scopre che a ucciderlo sono stati la moglie, dunque la sua stessa madre, Clitennestra, e il compagno di questa, Egisto. Clitennestra ed Egisto hanno ucciso Agamennone per vendetta. Clitennestra per vendicare il sacrificio della figlia Ifigenia, Egisto per vendicare suo padre, Tieste, a cui Atreo, padre di Agamennone, con l'inganno aveva servito da mangiare i figlioletti. Oreste vendica Agamennone uccidendo Egisto e Clitennestra, ma appena la vendetta è compiuta si materializzano le Erinni, divinità antiche che perseguitano chi si è macchiato di crimini violenti. Oreste è perseguitato dalle Erinni e troverà salvezza solo una volta giunto ad Atene. Là Atena stabilisce un metodo per porre fine alla vicenda: un processo, nel quale ciascuna delle parti esporrà le sue ragioni e in cui i cittadini voteranno per la condanna o l'assoluzione di Oreste. Atena partecipa alla votazione e vota in favore di Oreste. I voti sono pari e il giudizio viene risolto sulla base di una regola procedurale introdotta da Atena per cui, in caso di pareggio, Oreste sarà assolto. Oreste viene così assolto e le Erinni si trasformano nelle Eumenidi.

L'idea alla base della tragedia è che il sistema morale arcaico sia qualcosa che non è fatto per funzionare bene; la morale porta a dei vicoli ciechi che non possono essere oltrepassati con risorse rinvenibili nella morale stessa.

Un primo esempio a favore di questa conclusione è quello delle norme ricorsive. I personaggi della tragedia sono soggetti a una norma sanzionatrice dalla struttura ricorsiva, ossia una norma che obbliga a sanzionare un certo illecito, ma per la quale l'applicazione della sanzione costituisce a sua volta un illecito, che è presupposto della riapplicazione della norma medesima. Si tratta della norma che impone di vendicare la morte dei propri familiari. Dato quello che Agamennone ha fatto loro, Clitennestra ed Egisto hanno l'obbligo, derivante dalle leggi del sangue, di uccidere Agamennone, ma, obbedendo all'obbligo, i due commettono un illecito che obbliga, secondo le stesse leggi del sangue, il figlio dell'ucciso, Oreste, a sanzionare la condotta di Clitennestra ed Egisto. Ma anche la condotta di Oreste costituisce un illecito che dovrà essere sanzionato, e così via. Si crea un circolo vizioso all'interno della morale che andrà avanti all'infinito: la morale mostra un volto paradossale³.

Il secondo esempio è dato dal carattere dilemmatico della scelta di Oreste. Non solo Oreste per vendicare il padre compirà un illecito, ma lo stesso obbligo di vendicare il padre è incompatibile con un altro obbligo che pure pende sul suo capo, quello di rispettare sua madre. La tragicità della situazione di Oreste deriva dalla circostanza che giusto giusto l'assassino del padre sul quale dovrebbe riversarsi la

² Celano 2021, 77. Raz 2009.

³ Celano 2021, 343-346.

sua vendetta è la sua stessa madre. Su Oreste gravano due obblighi che, date le contingenze del caso, sono in conflitto – sull'agente gravano due obblighi contrari (l'obbligo di compiere e di non compiere la medesima azione) – e il cui conflitto non ha soluzione morale: Oreste sbaglia qualsiasi scelta prenda⁴.

La struttura ricorsiva di certe norme morali e le situazioni dilemmatiche che possono prodursi sono proprie di una concezione “pluralista” della morale. «Esiste una pluralità irriducibile di valori (...) incommensurabili, potenzialmente in conflitto (o in competizione) fra loro (...) L'etica è dunque, necessariamente, indeterminata, incompleta, e antinomica»⁵. Tutto ciò rende la morale qualcosa di impraticabile⁶, e questa impraticabilità della morale permette di intendere il rapporto tra diritto e morale in un nuovo modo. Diventa legittimo chiedersi se il diritto positivo – simboleggiato dal processo ad Atene – possa essere un modo per sanare le crepe della sfera morale⁷. L'Oresteia è dunque una sorta di mito sul passaggio dal mondo arcaico a un mondo urbano, civile, abituato alle procedure, da noi avvertito come più vicino, e sul rapporto – o meglio su un rapporto – tra il diritto naturale, già presente nel mondo arcaico, e il diritto positivo, appannaggio del mondo “civile”⁸.

Celano non sembra avere fiducia che il diritto possa dare soluzione alla crisi della morale. L'aspetto forse più interessante di questo viaggio letterario nel ragionamento pratico attiene proprio al fatto che Oreste, Eschilo e il pubblico aspettano con ansia di trovare nel diritto una non-soluzione al problema, e questa ricerca sembra essere del tutto sensata (finché Celano non ce la fa vedere per quello che è). Arrivati al momento del giudizio, le ragioni a favore e contro Oreste si mescolano nella mente dei giudici per produrre il verdetto in quello che assomiglia a un processo alchemico (la mente umana come “scatola nera”):

C'è una sorta di opacità della deliberazione, non sappiamo come ciascuno dei giudici pervenga al proprio giudizio; la deliberazione è qualcosa di privato. (...) Questo silenzio è uno dei cardini del congegno prospettato da Eschilo tramite Atena: la giurisdizione come metodo democratico assume il valore di una risoluzione e un superamento della difficoltà nella quale ci siamo trovati con Oreste proprio grazie a questa oscurità, proprio perché non sappiamo nulla. Ciascuno dei giudici decide per sé, e il modo in cui ciascuno perviene alla propria convinzione e al proprio voto resta soltanto suo.⁹

Una lettura estraniante del ragionamento pratico e del ruolo del diritto: «Eschilo fa questo gioco di prestigio, e noi ci sentiamo alleggeriti»¹⁰.

⁴ Celano 2021, 346-355.

⁵ Celano 2005, 175.

⁶ Celano 2021, 354.

⁷ Celano 2021, 354-355.

⁸ Celano 2021, 335, 360.

⁹ Celano 2021, 388.

¹⁰ Celano 2021, 392.

La morale e in generale la normatività umana sono – nelle parole di Celano – impraticabili. Questa idea, l'idea che la morale e la normatività siano “fatte male”, costituisce il tema del presente lavoro.

2. IL PROBLEMA DEI PARADOSSI MORALI NEL DIBATTITO RECENTE

Belief in God, or in many gods, prevented the free development of moral reasoning. Disbelief in God, openly admitted by a majority, is a very recent event, not yet completed. Because this event is so recent, Non-Religious Ethics is at a very early stage. We cannot yet predict whether, as in Mathematics, we will all reach agreement. Since we cannot know how Ethics will develop, it is not irrational to have high hopes.¹¹

Con queste parole si chiude *Reasons and Persons*, quella che è normalmente considerata l'opera più importante di Derek Parfit, nonché uno dei lavori più intensi e fecondi nell'ambito della filosofia morale contemporanea¹².

Poco meno di trent'anni dopo Jeff McMahan, filosofo in grande sintonia con l'approccio di Parfit, conclude un suo lavoro maturo con un'osservazione eloquente. McMahan si sta confrontando con un problema dell'etica della popolazione, la cui grammatica è stata una delle più originali creazioni di Parfit. La questione che sta affrontando è se sia possibile beneficiare una persona dandole la vita, se si diano benefici e danni esistenziali, e che ragioni per agire promanino da questi. Dopo aver esaminato con meticolosità varie posizioni possibili mettendone a fuoco le implicazioni controintuitive che queste generano, McMahan osserva:

This is a depressing point on which to conclude (...) Problems in the morality of causing people to exist seem to me the most difficult and intractable of all the problems of which I am aware in normative and practical ethics. They suggest that it is a real possibility that any moral theory that is both complete and coherent will have implications that are intuitively intolerable. It is these problems, therefore, rather than arguments in metaethics about the queerness of objective values, the connections between normativity and motivation, and so on, that seem to me to pose the greatest challenge to realism in ethics.¹³

Da un lato la speranza nella possibilità di inaugurare una nuova epoca per la teoria morale, dall'altro lo scoraggiamento derivante dalla sensazione che il progetto sia già fallito. E bisogna notare come il problema che McMahan evoca non riguardi solo l'etica della popolazione. Piuttosto, il fallimento dell'etica della popolazione sarebbe il contesto di scoperta del possibile fallimento del progetto di edificare una teoria morale: la mortificazione degli “*high hopes*” di Parfit.

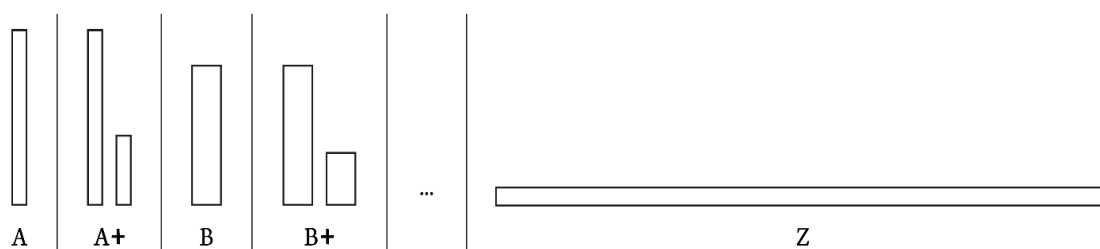
¹¹ Parfit 1984, 454.

¹² Come opere in italiano si considerino Pontara 2021 e Menga 2021

¹³ McMahan 2013, 34.

Come passiamo dalle speranze del passo di Parfit allo scoraggiamento di quello di McMahan? Cosa è successo nel mezzo? Ho accostato questi due passi scritti a distanza di trent'anni per ragioni espositive, ma la realtà è che Parfit aveva creato un quadro che già da solo avrebbe potuto motivare lo scoraggiamento di McMahan. In *Reasons and Persons* Parfit sembra dettare l'agenda di una scienza normale dell'etica, e, mentre lo fa, mette a fuoco alcuni problemi che riguardano le nostre intuizioni morali. In un certo qual modo l'agenda della scienza normale di Parfit consiste anzi proprio nella soluzione di questi problemi. Tuttavia, rispetto ai problemi delineati da Parfit potrebbe dirsi che poco o nessun passo avanti è stato fatto, e se un lavoro routinario di sviluppo del paradigma vi è stato, è più facile vederlo nel rafforzamento dei problemi originariamente configurati piuttosto che nel loro superamento¹⁴.

Di che natura sono i problemi assiologici individuati da Parfit? Si tratta di paradossi morali. Il più celebre è probabilmente il Paradosso della Mera Addizione.



La base di ciascuna barra del diagramma rappresenta il numero di persone esistenti, mentre l'altezza rappresenta la qualità della vita. A è una popolazione di miliardi di persone con una qualità della vita molto alta – possiamo immaginarla molto più alta di quella di qualsiasi paese contemporaneo. Z è una popolazione molto più vasta di A, in cui però la vita è appena degna di essere vissuta¹⁵. Z contiene più benessere di A – un esercito di bicchieri contenenti una sola goccia d'acqua può contenere più acqua di una brocca colma fino all'orlo – ma, per le intuizioni morali di molti, Z è peggiore di A. Che Z sia migliore di A perché contiene più benessere è una Conclusione Ripugnante: il *Rigetto della Conclusione Ripugnante* sembra essere una delle condizioni minime di accettabilità di una teoria morale.

A+ sembra essere almeno non peggiore di A. In A+ vi è un gruppo di persone grande quanto la popolazione complessiva di A che ha la stessa qualità della vita delle persone che esistono in A, e un gruppo aggiuntivo di persone che hanno una qualità della vita notevolmente inferiore a quella di A, ma che hanno comunque vite degne di essere vissute¹⁶. L'ineguaglianza che c'è in A+ e che non c'è in A potrebbe essere una ragione per ritenere che A sia preferibile ad A+. Tuttavia,

¹⁴ Si tratta questa di una possibile lettura del dibattito. Vi sono conclusioni molto diverse in proposito: eloquente un recente *position paper* che esprime ottimismo, Zuber et al. 2021.

¹⁵ Parfit 1984, 388.

¹⁶ Parfit 1984, 420.

l'eguaglianza che sussiste in A+ è frutto di una “mera addizione” di vite degne di essere vissute¹⁷. Se non riteniamo che la creazione di vite degne di essere vissute sia un male intrinseco, è difficile dare conto del perché la mera addizione che produce A+ renderebbe A+ peggiore di A. Chiamiamo “*Principio di Mera Addizione*” il principio secondo il quale l'addizione di vite degne di essere vissute non può diminuire il valore assiologico di una popolazione: il rispetto del Principio di Mera Addizione è una condizione di accettabilità di una teoria morale.

B sembra essere migliore di A+¹⁸. In B esistono tante persone quante ne esistono in A+. Inoltre la quantità totale di benessere è maggiore, e il benessere in B è suddiviso in modo più egualitario di quanto non lo sia in A+. Sembrerebbe che per preferire A+ a B non solo non dovremmo assegnare valore all'eguaglianza, ma dovremmo assegnargli un disvalore. Questo è assurdo: il *Principio di Non-Antieguagliantismo* è pure una delle condizioni minime di accettabilità di una teoria morale¹⁹.

Se B è meglio di A+, e A+ è almeno non peggiore di A, sembrerebbe doversi concludere che B sia migliore di A, o almeno non peggiore di A²⁰. La popolazione con benessere medio inferiore è preferibile o non è peggiore di quella con benessere medio maggiore. Se questa non fosse già una conclusione sufficientemente indigesta, possiamo ripetere il processo di mera addizione e redistribuzione fino a giungere alla popolazione Z. Come B è migliore o non peggiore di A, così Z sarà migliore o non peggiore di A. Tuttavia, sostenere questo vuol dire disconoscere una delle condizioni di accettabilità che ci eravamo posti, ossia il Rigetto della Conclusione Ripugnante.

Abbiamo allora tre giudizi che implicano una contraddizione. Se teniamo fermi il Principio di Mera Addizione e il Principio di Non-Antieguagliantismo dobbiamo concludere che Z sia almeno non peggiore di A. Ma allo stesso tempo riteniamo che Z sia peggiore di A (Rigetto della Conclusione Ripugnante). Chiamiamo “*Paradosso della Mera Addizione*” la contraddizione che è generata dai nostri giudizi²¹. Qualche giudizio deve cedere, ma quale?²²

Il problema dei paradossi morali è diverso dal problema del conflitto tra valori che deriva dall'adozione di una morale pluralista visto precedentemente. Su questo punto Parfit è esplicito. Se esiste una pluralità di valori, allora i valori possono entrare in conflitto. Se, ad esempio, crediamo che sarebbe contemporaneamente meglio che vi fosse più eguaglianza e più benessere, siamo costretti ad accettare l'eventualità in cui i nostri principi entrino in conflitto, perché può darsi il caso che una

¹⁷ Parfit 1984, 425.

¹⁸ Parfit 1984, 425-426.

¹⁹ Le espressioni “Principio della Mera Addizione” e “Principio di Non-Antieguagliantismo” sono prese da Ng 1989.

²⁰ Sulla distinzione tra la relazione “*not worse than*” e “*at least as good as*” si veda Parfit 1984, 430 ss.

²¹ Parfit 1984, 426.

²² Parfit 1984, 427; Parfit 2004 (1986), 13.

certa azione incrementi l'eguaglianza a scapito del benessere aggregato o viceversa. In questo caso, secondo Parfit, non siamo però davanti a una contraddizione. Abbiamo vari valori che rendono *prima facie* preferibile una certa azione o un certo stato di cose, e dobbiamo sforzarci di individuare quale sia l'azione o stato di cose *tutto considerato* preferibile²³. Per contro, lo schema dei paradossi morali nasce dalla circostanza che secondo le nostre intuizioni morali certe azioni o stati di cose sono *tutto considerato* migliori, peggiori, o uguali rispetto ad altre, epperò queste relazioni assiologiche tutto considerato sono tra di loro incoerenti. Il problema non è, ad esempio, che B sia migliore di A⁺ rispetto al benessere totale e all'eguaglianza, ma sia peggiore di A⁺ rispetto all'ideale del perfezionismo, bensì che B sembra essere migliore di A⁺ tutto considerato, ma questo giudizio tutto considerato non si sposa con i nostri altri giudizi tutto considerato che riguardano il rapporto tra A e B/Z e tra A e A⁺²⁴.

Nel precedente paragrafo abbiamo visto come la morale ospiti valori in conflitto che impediscono l'agglutinarsi di giudizi morali tutto considerato. In questo paragrafo abbiamo visto come, oltre a ospitare valori confliggenti, ospita anche giudizi tutto considerato contraddittori. La crisi della morale si è approfondita.

Il problema dei paradossi morali come problema della contraddittorietà è stato poi sviluppato oltre i confini delle originarie formulazioni parfitiane. Accanto alle riflessioni già citate di McMahan, in questo senso sono particolarmente rilevanti gli scritti di Gustaf Arrhenius e Larry Temkin²⁵.

I giudizi morali contraddittori che fungono da premesse del Paradosso della Mera Addizione possono essere ulteriormente indeboliti in modo da rendere ancora più forte la contraddizione. Questa via è stata percorsa da Arrhenius che, sulle orme del teorema dell'impossibilità di Arrow, ha elaborato diversi teoremi dell'impossibilità assiologica²⁶.

Temkin, invece, ha elaborato uno schema di paradosso che prende il nome di "*spectrum argument*". Dato un certo livello di dolore è possibile immaginare un decremento *marginale* di intensità accompagnato da un certo prolungamento del dolore tale per cui il dolore meno intenso ma più lungo sia peggiore di quello più intenso ma più breve. Così un giorno delle più estenuanti torture sarà migliore di due giorni di torture appena meno estenuanti; i due giorni saranno meglio di tre giorni di torture appena meno intense, e così via. Tutto ciò sembra essere abbastanza intuitivo. Per transitività, poi, la prima opzione - un giorno di tortura - sarà migliore dei tre giorni di torture meno atroci, e in generale delle opzioni in cui il dolore è progressivamente diminuito in modo marginale ma dura di più. Se però

²³ Il fatto che i valori siano plurali e che si diano situazioni tragiche complica o rende impossibile in certi casi stabilire cosa sia meglio tutto considerato (Parfit non si sofferma su questo aspetto, forse sottostimando il problema del conflitto tra valori).

²⁴ Parfit 1984, 427.

²⁵ Seguo qua Cowie 2020a, 4 ss.; e Cowie 2023, 465 ss.

²⁶ Arrhenius 2000.

ripetiamo questo processo più volte arriveremo alla conclusione secondo la quale il più leggero dei dolori immaginabili (come una puntura di zanzara) protratto per una certa quantità di tempo può essere peggiore di un giorno di atroci torture. Questa conclusione appare controintuitiva. Se consideriamo un decremento sufficientemente *cospicuo* nell'intensità di un dolore, non vi è incremento della durata che possa rendere il dolore meno intenso ma più lungo peggiore del dolore più intenso ma più breve. Siamo così arrivati in modo molto veloce a mettere a fuoco una contraddizione tra i nostri giudizi: (i) che per ogni decremento marginale nell'intensità del dolore vi è un incremento nella durata che rende il dolore più lungo e meno intenso peggiore; (ii) che per ogni decremento sufficientemente cospicuo nell'intensità di dolore non vi è un incremento nella durata che rende il dolore più lungo e meno intenso peggiore; (iii) che la relazione "migliore di\peggiore di" è una relazione transitiva²⁷.

Infine, va notato come, nel dibattito recente, il problema dei paradossi morali non sia nemmeno confinato all'ambito dell'etica della popolazione. Si tratta di un problema che emerge ad esempio nell'ambito del celeberrimo dibattito sul *Trolley Problem*²⁸. Se abbiamo ragioni morali per dirottare un vagone fuori controllo in modo che uccida una persona anziché cinque, perché mai avremmo ragioni morali per non spingere sul binario un grosso passante che si sta sporgendo dal cavalcavia per vedere la scena, in modo tale che il suo corpo blocchi la corsa del vagone fuori controllo salvando i cinque? Tradizionalmente l'indagine teorica si è concentrata sulla ricerca di un principio che giustifichi le nostre diverse risposte a casi del genere. Trovare la soluzione si è dimostrato notoriamente difficile. Con un certo sforzo – teorico o empirico – è possibile trovare dei principi che siano capaci di "contenere" le nostre risposte intuitive. Tuttavia, questi principi non sono in sé intuitivi. Si tratta, in effetti, di principi apparentemente privi di qualsiasi rilevanza morale. Nella misura in cui abbiamo ragione di credere che questi principi effettivamente causino i nostri giudizi intuitivi, possiamo dire di aver trovato dei principi che spiegano i nostri giudizi intuitivi, ma una spiegazione è cosa diversa da una giustificazione.

Si configura, nell'ambito del *Trolley Problem* il seguente dilemma per le nostre intuizioni. Possiamo optare per un principio che sia in sé intuitivo – che sembri qualcosa che abbia a che fare con la giustizia – ma che implica alcune conclusioni controintuitive: che ci impone di accettare alcuni giudizi contro-intuitivi. In alternativa, possiamo adottare un principio che collaziona le nostre risposte intuitive, il quale però non ha in alcun modo le sembianze di un principio morale, e pertanto è in sé contro-intuitivo. In un modo o nell'altro non ci resta che riformare il senso comune, accettando conclusioni contro-intuitive²⁹.

²⁷ Temkin 2012; Cowie 2023, 465-467.

²⁸ Foot 1967, Thomson 1985.

²⁹ Kagan 2016; Cowie 2023, 462-465.

3. MORAL ERROR-THEORY

Secondo il problema dei paradossi morali, messo in luce nel precedente paragrafo, le nostre intuizioni morali sono contraddittorie. La coperta delle intuizioni morali, delle soluzioni “familiari” ai problemi pratici, è troppo corta. Cosa ne segue?

Torniamo brevemente a McMahan. Ogni teoria morale ha implicazioni intuitivamente intollerabili, dunque

it is a real possibility that any moral theory that is both complete and coherent will have implications that are intuitively intolerable. It is these problems, therefore, rather than arguments in metaethics about the queerness of objective values, the connections between normativity and motivation, and so on, that seem to me to pose the greatest challenge to realism in ethics.

Secondo il “realismo morale” (i) i giudizi morali asseriscono fatti, e pertanto sono veri nella misura in cui asseriscono fatti veri, (ii) sussistono dei giudizi morali veri³⁰. L’idea di McMahan sembra essere che se non possiamo trovare soluzioni intuitive ai dilemmi morali allora il realismo morale deve essere scartato³¹.

In questa idea vi è un dettaglio interessante: la falsificazione del realismo procede da una fonte inaspettata.

Il realismo morale è una posizione metaetica e gli attacchi tipici alla posizione realista hanno solitamente pure carattere metaetico. Argomenti tipici contro il realismo sono – seguendo McMahan – l’idea per cui i valori oggettivi sarebbero delle entità bizzarre (*queer*), oppure il problema del legame tra valore morale e motivazione all’azione³². Il problema dei paradossi sembra essere strutturalmente diverso. Si tratta di un problema inerente all’etica sostanziale – non abbiamo una guida pratica per via della contraddittorietà delle nostre intuizioni – che tracima in ambito metaetico. La teoria morale potrebbe avere davanti a sé un destino radioso, come lo ha avuto la matematica. La capacità di svilupparsi, di sofisticarsi, sarebbe il segno o l’essenza dell’oggettività, della “realità”, dei valori morali. Ma se questa capacità di svilupparsi difetta, perché le nostre intuizioni morali sono contraddittorie, allora non abbiamo motivo di credere che esista una realtà morale.

Che io sappia, nella filosofia morale analitica contemporanea non vi sono stati tentativi di articolare questo problema per il realismo in modo sistematico. Non almeno fino a tempi recenti. Christopher Cowie ha colmato questa lacuna.

La mossa di Cowie è innovativa rispetto al panorama della filosofia morale analitica contemporanea, ma ha un precedente illustre nella filosofia morale vittoriana.

³⁰ Sayre McCord 2015.

³¹ In Celano l’idea per cui la morale sia incompleta e richieda il completamento del diritto si inserisce in una cornice metaetica realista (Celano 2005, 174 ss.). Non discuterò però la posizione metaetica di Celano.

³² Mackie 1977; Smith 1994.

Come osserva Cowie stesso, Sidgwick, nella prima edizione di *The Methods of Ethics*, aveva concluso che, se il problema della contraddittorietà dei giudizi etici non fosse stato risolvibile, tutto il sistema etico si sarebbe disintegrato.

Come noto, nel caso di Sidgwick, la contraddittorietà era dovuta al fatto che sia l'utilitarismo sia l'egoismo potrebbero costituire dei modelli normativi del ragionamento pratico razionali³³. Cowie ritiene che la contraddittorietà dei giudizi messa in luce dal problema dei paradossi morali ha conseguenze di larga portata per la teoria morale, al pari dell'impossibilità di risolvere razionalmente il conflitto tra egoismo e utilitarismo.

Così ragiona. Se uno dei giudizi o principi morali che compongono i paradossi è falso, allora tutti i nostri giudizi o principi morali sono falsi. Se abbiamo un insieme di giudizi o principi contraddittori, almeno uno dei giudizi che danno vita alla contraddizione deve essere falso. I paradossi morali individuati dalla recente filosofia morale sono insiemi di giudizi tra loro contraddittori. Dunque, almeno uno dei giudizi o principi morali che compongono i paradossi deve essere falso. Allora, tutti i nostri giudizi morali sono falsi.

A questo argomento Cowie, poi, imprime un ulteriore sviluppo. Siccome i giudizi morali sono asserzioni che mirano a riflettere delle verità morali, e non esistono verità morali, allora quando esprimiamo un giudizio morale stiamo commettendo un errore. Nel complesso l'argomento di Cowie utilizza la sussistenza dei paradossi morali per concludere in favore della teoria dell'errore ("*moral error-theory*")³⁴, posizione secondo la quale i giudizi morali sono falsi perché provano a riflettere delle verità morali, ma le verità morali non esistono.

Il passaggio più delicato dell'argomento consiste nel mostrare come, se uno dei giudizi o principi che compongono i paradossi è falso, allora tutti i nostri giudizi sarebbero falsi. Non è necessario ripercorrere per intero la sua argomentazione. Semplificando, l'idea generale di Cowie è che i giudizi o principi dei paradossi sono autoevidenti, e questa autoevidenza può essere compresa attribuendo loro lo status di verità concettuali: chiunque padroneggi il concetto di "bene" ("*goodness*") non può che riconoscere la verità di questi giudizi o principi³⁵. Si tratta, in altri termini, di truismi morali di cui non si può fare a meno se si vuole mantenere in piedi l'intero edificio della teoria pratica.

L'argomento di Cowie è un argomento in favore della teoria dell'errore. Tuttavia, come si diceva prima, è molto diverso dal tipo di argomenti che vengono presentati a supporto della teoria dell'errore. La tipica argomentazione che viene adoperata per difendere la teoria dell'errore consiste nel mostrare come i fatti morali sono delle entità in qualche modo bizzarre, *queer*: bizzarra sarebbe la modalità attraverso la quale dovremmo riuscire a entrare in contatto con questi fatti (piano

³³ Cowie 2020a, 1; Sidgwick 1874.

³⁴ Cowie 2020a, 6.

³⁵ Cowie 2020a, 7, 12.

epistemologico); bizzarro sarebbe il modo in base al quale i supposti fatti morali riuscirebbero a influenzare la nostra condotta (tesi motivazionale); bizzarro sarebbe il modo in cui i fatti morali dipenderebbero dai fatti naturali (problema concernente la sopravvenienza); e bizzarro, da un punto di vista metafisico, sarebbe il carattere irriducibilmente normativo dei fatti morali³⁶.

L'argomento a favore della teoria dell'errore che muove dal problema dei paradossi presenta varie differenze rispetto all'argomento classico della bizzarria, e, pur non escludendo l'impiego dell'argomento della bizzarria, potrebbe permettere di superarne alcune criticità.

In primo luogo, e come già anticipato, la crisi del ragionamento pratico su cui viene posta l'enfasi ha carattere per così dire "interno" e non invece "esterno" come per l'argomento della bizzarria. In generale l'argomento della bizzarria riconduce la crisi della morale a un problema esterno alla morale, ossia a un insieme di domande che non sono esse stesse domande morali, bensì domande *sulla* pratica, l'idea, della morale. Uno scetticismo "archimedeo" che sta sopra la morale e la giudica dall'esterno³⁷. Secondo un certo modo di vedere – che si ritrova nell'opera tarda di Ronald Dworkin – non è possibile contestare le conclusioni morali sulla base di preoccupazioni di carattere metafisico³⁸. Al contrario, il problema dei paradossi morali è un problema che sorge proprio dall'interno del ragionamento morale. Non è casuale che siano proprio i filosofi morali che si sono occupati di etica sostanziale e non di metaetica ad avere manifestato inquietudine in proposito (si pensi ancora una volta a McMahan). In questo senso il problema dei paradossi costituisce un problema per il realismo, anche se accettiamo l'idea di Dworkin per cui non sia possibile contestare conclusioni morali se non su basi a loro volta morali³⁹.

In secondo luogo, l'argomento dei paradossi permette di attenuare lo svantaggio in termini di intuitività della teoria dell'errore morale. Anche la teoria dell'errore potrebbe avere esiti contro-intuitivi. Consideriamo giudizi come "torturare i bambini è sbagliato", oppure "L'Olocausto è qualcosa di moralmente problematico". Secondo la teoria dell'errore sono giudizi falsi⁴⁰. Siccome la teoria dell'errore

36 Nell'analisi dell'argomento seguo Olson 2018, 62-63; l'argomento della queerness è stato elaborato nella famosa opera di Mackie 1977.

37 Dworkin 2011, 30-32.

38 Dworkin 2011, 25.

39 Cowie 2020a, 2.

40 Cowie 2023, 457-458. Non bisogna comunque imputare più contro-intuitività di quella necessaria. Per la teoria dell'errore morale questi giudizi sono falsi, ma questo non vuol dire che i giudizi opposti – contraddittori – siano veri. Non vuol dire che sia vero che "torturare i bambini sia permesso" o che "l'Olocausto sia qualcosa di moralmente non-problematico". Anche questi giudizi come i precedenti due sono falsi. Ciò non vuol dire però che la teoria dell'errore morale non resti contro-intuitiva nel suo negare che l'Olocausto sia moralmente problematico, seppure in questo senso meno impegnativo Cowie 2023, 460-462. In aggiunta, bisogna ricordare che il carattere contro-intuitivo può essere ulteriormente ridotto adottando una posizione finzionalista come fa, ad esempio, Joyce, uno dei principali esponenti della teoria dell'errore (Joyce 2001).

morale implica conclusioni contro-intuitive, l'argomento della bizzarria è spesso complementato con altri argomenti che servono a dar conto del carattere contro-intuitivo della teoria dell'errore: argomenti che mostrano perché – su quali basi evolutive e con quale funzione sociale – tendiamo a scambiare le nostre reazioni emotive contro certe condotte con la percezione di proprietà reali (*mind-independent*) e irriducibilmente normative⁴¹. Quale sia l'efficacia di questi argomenti non è quello che importa. Basta osservare come il problema dei paradossi morali apra una diversa via per ovviare al problema della contro-intuitività della teoria dell'errore.

Su questo sfondo, l'idea generale dell'argomento dei paradossi è che, se anche l'accettazione della teoria dell'errore avesse delle conseguenze contro-intuitive sul piano morale – vedi i giudizi sul torturare i bambini o sull'Olocausto –, questo non sarebbe un difetto decisivo, perché, alla fine, anche le altre posizioni metaetiche avrebbero implicazioni morali pesantemente contro-intuitive⁴². Infatti, il rinvenimento dei paradossi mostra come non è possibile trovare una teoria che rispetti tutti i nostri giudizi intuitivi, dal momento che le nostre intuizioni sono contraddittorie; il realismo non può tollerare la contraddittorietà dei giudizi⁴³ e dunque, per risolvere il conflitto, diventa necessario accettare una conclusione contro-intuitiva – se il realismo morale fosse vero, dovremmo accettare dei giudizi contro-intuitivi; dal canto suo la teoria dell'errore nega che si debba accettare come vero qualsivoglia giudizio morale, e dunque impedisce che nascano delle contraddizioni tra i giudizi; dunque la teoria dell'errore e la teoria dell'errore soltanto evita che debbano accettarsi delle conclusioni morali contro-intuitive⁴⁴.

Dato il paradosso dell'argomento a spettro elaborato da Temkin, se il realismo fosse vero dovremmo accettare che uno dei nostri giudizi morali rilevanti sia falso. Mettiamo caso che sia falso il giudizio per cui “se consideriamo un decremento sufficientemente cospicuo nell'intensità di un dolore, non vi è incremento nella durata che possa rendere il dolore meno intenso ma più lungo peggiore del dolore più intenso ma più breve”. Se questo principio è falso, dovrebbe essere vero il giudizio contrario: per qualsiasi decremento nell'intensità di dolore, vi è un incremento della sua durata tale per cui il dolore meno intenso ma più lungo è peggiore. Ossia, un giorno delle più efferate torture è peggiore di un leggero prurito che dura per tutta la vita. Il realismo dovrebbe implicare che giudizi del genere sono veri. Invece la teoria dell'errore morale (e solo essa) eviterebbe che si debbano accettare conclusioni controintuitive del genere. Presupposto di queste conclusioni, infatti, è che

41 Olson 2018, 66. Un autore che ha ampiamente utilizzato questa strategia è stato Joyce (Joyce 2001).

42 Cowie 2023, 459-460.

43 Questa sembra essere una premessa tacita dell'argomento di Cowie. Da notare come sia in contrasto con la posizione di Celano che qualifica la sua concezione pluralista del diritto naturale come “realismo etico politeista” (Celano 2005, 174).

44 Cowie 2023, 462.

vi siano delle verità morali, ma l'esistenza di queste ultime è negata dalla teoria dell'errore.

Ancora, se il realismo fosse vero dovremmo rifiutare qualcuno dei principi morali la cui contraddittorietà genera il Paradosso della Mera Addizione. Dovremo ritenere falso e dunque rifiutare o il Principio di Non-Antiegalitarismo, o il Principio della Mera Addizione, oppure il Rigetto della Conclusione Ripugnante. In ogni caso, dovremmo ritenere vero e dunque accettare il principio opposto a quello rigettato, e questo vuol dire che dovremmo ritenere vero e accettare un principio che sia pesantemente contro-intuitivo. La teoria dell'errore morale, di nuovo, può evitare questa deriva.

Infine, se il realismo fosse vero, nel *Trolley Problem* dovremmo o ritenere falsi alcuni dei nostri giudizi intuitivi e dunque accettare dei giudizi pesantemente contro-intuitivi, oppure ritenere veri solo i giudizi intuitivi che abbiamo, adottando un principio ombrello che in sé è contro-intuitivo. L'alternativa potrebbe essere, ad esempio, tra, da un lato, l'accettazione dell'Utilitarismo e del giudizio, da esso implicato, per cui sia giusto spingere una persona sotto un treno per salvare cinque persone, e, dall'altro lato, l'accettazione di una teoria che distingue tra vittime di condotte che richiedono l'impiego di forza fisica e vittime di condotte che non richiedono l'impiego di forza fisica (un fattore che, sebbene possa funzionare come spiegazione psicologica delle nostre reazioni, difficilmente sembrerebbe essere moralmente decisivo). Anche in questo caso la teoria dell'errore morale ci toglierebbe dai guai dal momento che non dovremmo accettare come vero né il giudizio per cui sia giusto spingere una persona sotto il treno, né un'improbabile teoria che assegni rilevanza decisiva all'impiego della forza muscolare. Questi giudizi pretendono di essere veri, ma secondo la teoria dell'errore morale non lo sono⁴⁵.

Tutto ciò non vuol dire che l'accettazione della teoria dell'errore abbia sotto il profilo morale implicazioni meno contro-intuitive del realismo o di altre posizioni metaetiche. Vuol dire però che la forza dell'argomento delle implicazioni contro-intuitive si smorza perché sotto il profilo della contro-intuitività morale la differenza tra la teoria dell'errore e le altre teorie metaetiche si è assottigliata.

Questo l'argomento di Cowie. Possono sollevarsi dubbi sulla forza di questo argomento. Anche solo limitatamente all'argomento a spettro, al Paradosso della Mera Addizione e al *Trolley Problem*, non è chiaro quale sia il vantaggio della teoria dell'errore rispetto al realismo sul fronte dell'intuitività dei giudizi. Del resto, per la teoria dell'errore a non essere veri non sono solo i giudizi contro-intuitivi, ma anche i giudizi *intuitivi: tutti i giudizi intuitivi*⁴⁶. Ad ogni modo oggetto del presente lavoro non è la critica all'argomento di Cowie, bensì la presentazione di un diverso uso teorico del problema dei paradossi.

45 Cowie 2023, 462-469.

46 Cowie prende in considerazione varie obiezioni (Cowie 2023, 473 ss.).

4. ACCETTARE CONCLUSIONI RIPUGNANTI

L'idea che vi sia un legame tra etica sostanziale e metaetica è affascinante⁴⁷. Non è detto però che il legame esatto, il legame più interessante o l'unico legame sia quello delineato da Cowie. In questo e nel prossimo paragrafo delineerò un itinerario alternativo, per così dire “rovesciato”, che collega etica normativa e metaetica. Non argomenterò se questo itinerario sia preferibile rispetto a quello tracciato da Cowie: per il momento mi sembra sufficiente provare ad abbozzare quale questa alternativa possa essere.

Per farlo è utile tornare al Paradosso della Mera Addizione e considerare una soluzione eccentrica. Non è forse un caso che lo stesso Cowie, prima di sviluppare il suo argomento che supporta la teoria dell'errore, abbia dedicato una breve monografia a questa posizione⁴⁸. Chiamiamo questa posizione “Accettazione della Conclusione Ripugnante”. Diversi autori hanno proposto di risolvere il Paradosso della Mera Addizione accettando la Conclusione Ripugnante⁴⁹. In questo contesto ci limiteremo ad analizzare l'argomento di Torbjörn Tännsjö.

L'idea fondamentale dell'argomento di Tännsjö è che le persone tendono a sovrastimare la qualità della loro vita e, correlativamente, a sottostimare le condizioni di una vita appena degna di essere vissuta. Spesso le vite delle persone fortunate del mondo contemporaneo sono solo appena degne di essere vissute. I più fortunati del pianeta vivono “schiacciati” attorno al livello Z, e sono ben lontani dall'approssimarsi alle alture del livello A. Ma se così è – continua l'argomento – allora la Conclusione Ripugnante *cessa* di essere ripugnante: non è ripugnante – non è intuitivamente erroneo – che a una popolazione con un'elevata qualità della vita vada preferita una popolazione molto più vasta composta da persone che hanno un livello di benessere pari a quello delle persone più fortunate del nostro mondo⁵⁰. Come Cowie osserva, siamo davanti a un *debunking argument*: nello specifico, un argomento che spiega come una conclusione che sembrava irrimediabilmente falsa in realtà sia vera⁵¹.

È molto importante, però, capire in che senso la Conclusione Ripugnante “cesserebbe” di essere ripugnante. I *debunking arguments* o argomenti genealogici pretendono di spiegare perché le nostre intuizioni morali sarebbero fallaci, ma questa

47 Cfr. Schroeder 2015

48 Cowie 2020b. Le idee illustrate in questo paragrafo e nei prossimi due costituiscono il nucleo della mia tesi di dottorato “Etica della popolazione. Sul ruolo delle intuizioni morali”, scritta e discussa (2019) prima della pubblicazione della monografia in cui Cowie analizza questa posizione “eccentrica”, e fuori da ogni contatto con l'opera di Cowie. Non è certo casuale che sia il lavoro di Cowie sia il mio lavoro – entrambi interessati al rapporto tra etica e metaetica – si fossero concentrati su questa posizione.

49 Si vedano Tännsjö 2004, Ryberg 2004 e Huemer 2008. Cowie 2020b si concentra anche sull'opera di Dasgupta e Wolf.

50 Tännsjö 2004, 223 ss.

51 Cowie 2020b, 37-38.

spiegazione difficilmente può avere l'effetto di modificare psicologicamente le nostre intuizioni. Questo è a mio avviso il caso dell'argomento di Tännsjö, che è incapace di modificare le nostre intuizioni e si risolve semmai nella *prescrizione* di trascurarle. Siamo sotto un incantesimo che non ci permette di vedere come stanno le cose, ma sapere che siamo sotto l'incantesimo, e come esso opera, non è sufficiente a rompere l'incantesimo facendoci riacquisire i nostri sensi. Se, in altri termini, riconosciamo di essere in preda a delle illusioni morali, queste illusioni sono illusioni persistenti. La Conclusione Ripugnante, dunque, non cessa affatto di *sembrarci* ripugnante, e continua invece a risultare altamente implausibile rispetto alle nostre intuizioni morali di primo livello⁵². Siamo davanti a un argomento genealogico psicologicamente impotente, che intende però destituire alcune intuizioni morali - una porzione della nostra psicologia - dal ruolo di parametro di correttezza.

Leggendo l'argomento di Tännsjö potrebbe sembrare diversamente. Potrebbe sembrare che in fondo l'argomento continui ad assegnare alle nostre intuizioni morali il ruolo di parametri di correttezza. Potrebbe dirsi: se le vite dei più fortunati del mondo contemporaneo vanno collocate più o meno al livello Z, e il livello A sarebbe raggiunto solo da vite inimmaginabili secondo i nostri standard, allora non *sembrerebbe* assurdo preferire Z ad A.

Ma, a guardare meglio, è chiaro che l'argomento di Tännsjö tolga alle intuizioni morali il ruolo di parametri di correttezza. Per Tännsjö le persone più fortunate della nostra epoca vivono nelle prossimità di Z, ma vivono nella prossimità di Z anche le persone meno fortunate, e quelle ancora meno fortunate, e così via. Non solo hanno vite "Z" gli svedesi ricchi e perfettamente sani, ma anche gli svedesi ricchi che hanno avuto un'incidente menomante, e anche gli svedesi che hanno avuto un incidente menomante e che non sono nemmeno tanto ricchi, e così via. Per questo, se anche fosse vero che accettare la Conclusione Ripugnante non farebbe torto alle nostre intuizioni nel momento in cui immaginiamo Z come il mondo degli svedesi ricchi e sani di oggi⁵³, la proposta di Tännsjö implicherebbe comunque l'accettazione di conclusioni contro-intuitive: semplificando, che una persona ricca e sana abbia la stessa qualità della vita di una persona decisamente meno abbiente e decisamente meno in salute. Del resto, per Tännsjö, a dispetto delle nostre fantasie, le vite che abbiamo non possono essere molto migliori di vite appena degne di essere vissute. Si tratta di una posizione che destituisce in modo esplicito le nostre intuizioni dal ruolo di specchio obbiettivo della nostra condizione.

Il fatto che l'argomento dell'Accettazione sia psicologicamente impotente e che si traduca nella destituzione delle nostre intuizioni morali è gravido di conseguenze.

⁵² Con ciò intendo intuizioni morali su cosa sia giusto o buono fare.

⁵³ Secondo alcuni critici a essere contrario alle nostre intuizioni non sarebbe tanto l'accettazione del valore del mondo Z, quanto il fatto che il mondo Z vada preferito al mondo A. Si vedano in proposito le posizioni di Holtug e Arrhenius analizzate dallo stesso Tännsjö (2004, 227).

Per farle emergere più chiaramente soffermiamoci su un problema specifico della posizione di Tännsjö, messo in luce da Arrhenius.

Come abbiamo visto sopra, alla Conclusione Ripugnante si può giungere su basi teoriche esili, ossia in base alla combinazione del Principio della Mera Addizione e del Principio di Non-Antiegalitarismo. Tuttavia, Tännsjö si occupa primariamente della Conclusione Ripugnante per difendere l'Utilitarismo edonista. Nello specifico Tännsjö ha in mente il Principio Totale in versione edonista, secondo il quale il migliore scenario è lo scenario che contiene la maggiore quantità totale di benessere, inteso come piacere⁵⁴. Un argomento contro il Principio Totale è che questo implica la Conclusione Ripugnante. Accettare la Conclusione Ripugnante è un modo per Tännsjö di rispondere a una critica che può essere mossa al Principio Totale⁵⁵. Ma l'adozione del Principio Totale, nota Arrhenius, peggiora la posizione di Tännsjö⁵⁶.

Se lo scenario preferibile è quello che contiene più benessere, allora, non solo dobbiamo preferire a un mondo pieno di persone con un elevato livello di benessere un mondo molto più popolato in cui ciascuno ha una vita appena degna di essere vissuta - Z migliore di A -, ma dobbiamo preferire anche un mondo molto più popolato in cui un enorme numero di persone vive una vita appena degna di essere vissuta, e in cui vi sono anche tantissime persone che vivono una vita degna di *non* essere vissuta, ossia una vita con benessere netto negativo - Z^* migliore di A . Sebbene le persone che vivono una vita degna di non essere vissuta contribuiscono in modo negativo al valore dello scenario, questo disvalore viene compensato dal valore dell'enorme numero di persone che vivono vite degne di essere vissute, e il valore positivo netto di queste può superare il valore delle vite con un elevato livello di benessere dell'altro scenario. Arrhenius chiama questa conclusione "la Conclusione Veramente Ripugnante" e ritiene che sia una conclusione difficile da accettare.

Come risponderebbe Tännsjö al problema sollevato da Arrhenius?⁵⁷ Tännsjö insisterebbe - secondo Cowie - nella difesa del suo argomento genealogico, ritenendo che le nostre intuizioni morali sono compromesse anche quando valutiamo il mondo Z^* : nemmeno la Conclusione Veramente Ripugnante sarebbe ripugnante dopotutto, e andrebbe accettata. Nello specifico l'argomento, ancora una volta, consisterebbe nel mostrare come le nostre intuizioni sul valore delle vite sono inaffidabili. Quando consideriamo Z^* siamo portati a immaginare un mondo infernale, ma Z^* può essere un mondo in cui coesistono due gruppi di persone, il primo che trascorre vite lunghe composte da anni che usualmente sono marginalmente positivi

⁵⁴ Parfit 1984, 387.

⁵⁵ Tännsjö 2004, 224.

⁵⁶ Arrhenius 2003.

⁵⁷ Secondo Cowie la via più promettente (ma nemmeno questa alla fine convincente) consiste nel negare che l'Accettazione della Conclusione Ripugnante implichi l'Accettazione della Conclusione Veramente Ripugnante (Cowie 2020b, 78 ss). Non sarebbe però questa la via seguita da Tännsjö.

ma a volte negativi, il secondo che trascorre vite lunghe composte da anni che usualmente sono marginalmente negativi e a volte sono positivi. Se intendiamo così Z^* , lo scenario cessa di essere ripugnante⁵⁸.

Contro questo argomento Cowie nota che l'Accettazione della Conclusione Veramente Ripugnante dipenderebbe da una particolare rappresentazione delle vite degne di non essere vissute di Z^* . Le vite degne di non essere vissute di Z^* potrebbero però essere diverse: potrebbero non essere vite appena degne di non essere vissute, ma vite piene di tormenti. Accettare che Z^* sia meglio di A anche in questi casi sembra davvero un rospo troppo grosso da ingoiare.

In questa versione del suo argomento Tännsjö contesterebbe le nostre intuizioni su cosa sia una vita degna di non essere vissuta. Rispetto a questa versione dell'argomento, mi sembra che la risposta di Cowie sia convincente. Ma possiamo sviluppare la posizione di Tännsjö in modo diverso. Invece che contestare le nostre intuizioni su cosa sia una vita degna di non essere vissuta, potremmo, ad esempio, bersagliare le nostre intuizioni sull'importanza del creare vite degne di essere vissute (anche se appena degne di essere vissute). Potremmo dire, in altri termini, che le nostre intuizioni morali sono difettose nel non cogliere l'importanza della creazione di vite aggiuntive, importanza che può bilanciare il disvalore insito nella creazione di vite degne di non essere vissute, sia che queste siano vite appena degne di non essere vissute, sia che siano vite tormentate.

Questa mossa mi sembra che sia coerente con lo spirito dell'Utilitarismo, con un approccio genealogico e, in generale, con un programma etico rivoluzionario che destituisca le intuizioni morali di primo livello, il senso comune, da parametro di valutazione delle varie teorie. Nel contesto della difesa dell'Utilitarismo questo approccio rivoluzionario è stato difeso, ad esempio, da Peter Singer e da Joshua Greene⁵⁹. Sebbene Greene non si occupi direttamente di etica della popolazione, quanto dei dilemmi legati al *Trolley problem*, è eloquente una sua osservazione (che prova, tra l'altro, la connessione tra i due dibattiti):

Our resistance to making more happy people is not, I think, a moral one. Rather, we living humans are engaged in a conspiracy against the most underrepresented people of all, an underrepresented majority, in fact: the helpless hypothetical masses who, thanks to our selfish choices, never get the chance to even protest their non-existence⁶⁰.

Secondo Greene le nostre intuizioni morali sarebbero profondamente difettose nella misura in cui non ci permettono di avvertire il valore insito nella creazione di nuove vite.

⁵⁸ Questo argomento non si ritrova nell'articolo di Tännsjö (sebbene ve ne siano tracce - si veda Tännsjö 2004, 224). Seguo l'analisi che Cowie fa di Tännsjö, che sembra essere basata su un manoscritto di quest'ultimo (Cowie 2020b, 75 ss.).

⁵⁹ Singer 2005; Greene 2013; si rinvia anche a Rocchè 2020 e 2024.

⁶⁰ Greene 2013, 326.

Una volta imboccata questa strada mi sembra inevitabile che il dibattito diventi un dialogo tra sordi. Una parte avanzerà un principio morale generale o una teoria. L'altra parte noterà come quel principio o teoria abbia un'implicazione pesantemente contro-intuitiva sotto il profilo morale. La prima parte non si scomporrà: l'implicazione è controintuitiva ma va accettata, perché le nostre intuizioni morali sono tutt'altro che infallibili (al contrario, sono pregiudizi, sono la fonte del problema). L'altra parte troverà allora un'ulteriore implicazione del principio o della teoria che cozza in modo ancora più deciso con le intuizioni di senso comune. Ma questa ulteriore dimostrazione di contro-intuitività sarà irrilevante o almeno non decisiva per la prima parte, dal momento che questa ha già messo in dubbio il valore delle intuizioni morali. Per quanto ingegnosi possano essere i controesempi forniti dal critico, questi hanno perso la capacità di confutare le sue proposte. E viceversa le sue proposte agli occhi del critico risulteranno inaccettabili e gratuite. Le due parti non stanno più giocando secondo le stesse regole. Il dibattito ha dato vita a due stili morali diversi.

5. STILI MORALI: IL METODO DELLA FAMILIARIZZAZIONE E IL METODO DELL'ACCETTAZIONE

Questi due stili non sono stati teorizzati⁶¹, e non vanno intesi come ipotesi ricostruttive sulle posizioni degli autori coinvolti nei dibattiti contemporanei di filosofia morale, bensì come “radicalizzazioni” dei loro argomenti (torneremo a breve su questo punto).

Possiamo chiamare i due stili “il Metodo intuizionistico”, o “il Metodo della familiarizzazione”, e “il Metodo dell'Accettazione”. Secondo il Metodo della familiarizzazione un giudizio morale per essere giustificato deve essere, in qualche modo, in accordo con le intuizioni morali dell'agente: deve *sembrare* giusto. Riflessione e dialogo morale preservano una funzione, dato questo ruolo chiave delle intuizioni di partenza? La riflessione e il dialogo morale funzionano come un contagio (endogeno o esogeno), nel senso che sono situazioni grazie alle quali può accadere che giudizi prima familiari cessino di essere familiari e che giudizi prima non-familiari (“inaccessibili” da un punto di vista intuitivo-emotivo) diventino familiari. Il rigetto o l'accoglimento di un certo giudizio si riduce alla registrazione delle nostre reazioni psicologiche. Per il Metodo dell'accettazione, invece, la familiarità di un giudizio non è condizione necessaria per la giustificazione del giudizio medesimo. Quello che sembra corretto potrebbe essere effettivamente sbagliato, e viceversa. Per questo il Metodo dell'Accettazione implica che in certi casi dobbiamo accettare conclusioni che sono contro-intuitive.

61 Sebbene vi siano operazioni teoriche che vanno in questa direzione, come la distinzione tra intuition chasing e bullet biting (Greene 2014 718-719).

Come detto, Cowie ha ben presente la posizione di Tännsjö⁶², e nei suoi scritti sulla teoria dell'errore morale non manca di considerare la posizione di chi rigetta una delle intuizioni che danno vita ai paradossi morali, ossia accetta conclusioni contro-intuitive⁶³. L'analisi di Cowie però a mio avviso non cattura l'essenza della posizione di chi argomenta a favore dell'accettazione di conclusioni contro-intuitive nel dibattito morale contemporaneo (penso a Tännsjö, Singer e Greene). Per Cowie la domanda fondamentale riguarda il se le conclusioni ripugnanti accettate sarebbero meno contro-intuitive dell'adozione della teoria dell'errore morale (se lo sono, abbiamo argomenti a favore del realismo, se non lo sono, abbiamo argomenti a favore della teoria dell'errore). Cowie riduce le ragioni a favore e contro una conclusione al loro carattere contro-intuitivo e procede a effettuare un bilanciamento di queste ragioni, che è un calcolo sulla contro-intuitività. Ma per i sostenitori dell'accettazione delle conclusioni contro-intuitive il carattere moralmente contro-intuitivo non è una ragione per abbandonare la loro teoria morale⁶⁴.

Quale può essere una ragione per adottare uno stile anziché un altro? Il mio suggerimento è che una ragione può provenire dalle posizioni metaetiche che vengono adottate. Invece che andare dall'etica sostanziale alla metaetica, si va dalla metaetica all'etica sostanziale.

Contrapponiamo due famiglie di posizioni in senso lato metaetiche: da un lato, il realismo e l'oggettivismo delle ragioni, dall'altro lato, il non-cognitivismo, la teoria dell'errore, e il soggettivismo delle ragioni nella sua variante internista. *Grosso modo* – sfumando una enormità di differenze, la cui considerazione sarebbe impossibile adesso – la faglia attorno alla quale queste due famiglie si dividono è costituita dal collegamento tra valori, ragioni, morali e il nostro apparato psicologico-emotivo-motivazionale, le nostre proattitudini. Per il non-cognitivismo, per la teoria dell'errore e per l'internismo delle ragioni, ossia per posizioni che ricomprenderò nel “nichilismo”, i nostri valori e ragioni finali dipendono dalle nostre proattitudini. Per la famiglia del realismo e dell'oggettivismo delle ragioni esistono invece delle verità morali, che, in quanto tali, sono indipendenti dalle proattitudini delle persone⁶⁵.

62 Cowie 2020.

63 Cowie 2022, 14 ss.

64 In un certo qual modo ogni teoria deve essere intuitiva per riscuotere l'assenso di persone in buona fede. Sarebbe ingiusto dire che Tännsjö, Singer e Greene non aspirino a incontrare le nostre intuizioni con le loro teorie. Il punto è che questi autori ritengono che una teoria morale non sia meno solida per il non appagare certe intuizioni morali. Stabilire quali intuizioni e di quale genere siano importanti e quali, invece, siano destituite di rilevanza è onere delle loro proposte teoriche. Un onere gravoso, ma in sé direi che questo atteggiamento selettivo rispetto alle intuizioni non è contraddittorio.

65 Sul collegamento tra teoria dell'errore e non-cognitivismo si veda Joyce 2001, 200 ss., e anche Parfit 2011b, 384 (che impiega anche il termine “nichilismo”). Sull'oggettivismo e il soggettivismo delle ragioni Parfit 2011a Cap. II, III. Sulla differenza tra ragioni esterne e ragioni interne Williams 1981 (1979). Sul come il soggettivismo internista (non-analitico) sia una forma di nichilismo,

Se si ritiene che esistano delle verità morali, e che i giudizi morali abbiano la funzione di riflettere queste verità, allora è possibile credere che certe nostre intuizioni morali, per quanto tenaci, siano radicalmente sbagliate⁶⁶, siano dei meri pregiudizi. Le intuizioni sono tenaci, ma lo sono solo da un punto di vista psicologico: un punto di vista irrilevante sul piano normativo. Se si ritiene che esistano verità morali e che le nostre intuizioni potrebbero essere radicalmente sbagliate, allora il metodo del ragionamento morale ammetterà che in certi casi si debbano accettare delle conclusioni contro-intuitive. Una prospettiva oggettivista, realista, rende sensata l'opzione in favore del Metodo dell'accettazione. Il Metodo dell'accettazione sarebbe invece privo di senso se si muove da una posizione morale nichilista, ossia una posizione che neghi l'esistenza di valori morali oggettivi, sia questa associata a una tesi non-cognitivista oppure alla teoria dell'errore morale.

Il Metodo della Familiarizzazione richiede simmetricamente l'adozione del nichilismo oppure è compatibile con il realismo? Per rispondere a questa domanda servirebbe addentrarsi nel problema molto di più di quanto non possa farsi adesso. In generale, mi sembra che se il Metodo della Familiarizzazione sia compatibile o no con il realismo dipenda dal se in esso vi sia spazio o meno per la nozione di errore morale. Se il Metodo della Familiarizzazione non lascia spazio all'idea per cui l'agente versi in uno stato di errore morale (da distinguere rispetto agli errori sui fatti fisici) non è chiaro – sebbene sia richiesta una riflessione ulteriore – come possa lasciare spazio all'idea di verità morale, e dunque essere compatibile con il realismo.

Ma in che senso la teoria dell'errore – il nichilismo – avrebbe un *metodo* del ragionamento suo caratteristico? Si potrebbe obiettare che il Metodo della Familiarizzazione non solo non esclude l'adozione di una forma di realismo, ma esclude in modo deciso proprio l'adozione della teoria dell'errore. Infatti, ho presentato il Metodo della Familiarizzazione come un metodo per stabilire quando un giudizio morale è *giustificato*. Ma per la teoria dell'errore morale nessun giudizio morale parrebbe poter essere giustificato, dal momento che non esistono verità morali cui tendere. Se il Metodo della Familiarizzazione è un metodo del ragionamento che stabilisce in modo astratto quali tipi di ragioni abbiamo, dovrebbe essere assolutamente incompatibile con la teoria dell'errore morale per la quale non abbiamo ragioni per agire (le asserzioni sulle ragioni per agire sono tutte indistintamente false)⁶⁷.

Credo che possiamo difendere il legame tra Metodo della Familiarizzazione e teoria dell'errore morale. Tizia darebbe entrambe le braccia per vedere i suoi figli crescere in un mondo democratico, libero e pacifico. Dobbiamo per forza

espressione con la quale Parfit designa la teoria dell'errore, Parfit 2011b, 273, 442 (su Williams come nichilista).

⁶⁶ Si veda come esempio di questa sensibilità l'analisi che Peter Unger fa dello schiavismo (Unger 1996, 14 ss.).

⁶⁷ Questa obiezione concederebbe però un legame tra Metodo della Familiarizzazione e soggettivismo delle ragioni in versione esternista.

intendere Tizia come una persona che professa una qualche forma di realismo, o può Tizia coerentemente adottare la teoria dell'errore? Direi che l'atteggiamento di Tizia è compatibile con il rigetto del realismo. Tizia ha dei valori per cui morire, e non crede che questi valori riflettano delle verità morali, ma che dipendano dalle sue proattitudini. Crede che, in modo del tutto contingente, lei abbia certe proattitudini, ma queste proattitudini sono la sua identità. Questa credenza nella contingenza dei valori non pregiudica forse l'adozione di un vero e proprio punto di vista morale? Non mi sembra. Sebbene creda che i suoi valori siano completamente contingenti rispetto alle sue proattitudini, Tizia concepisce non solo ragioni ipotetiche, ma anche ragioni categoriche. Poniamo che Tizia sappia che verosimilmente i suoi figli *non* cresceranno in un mondo democratico, libero e pacifico, e che possa, in qualche modo, eliminare questo suo desiderio che sarà irrimediabilmente frustrato. Eliminare un desiderio la cui frustrazione le causerebbe dolore è: una soluzione razionale, qualora si tratti di un mero desiderio. Questa opzione però non soddisfa Tizia, perché dato il carattere delle sue proattitudini – proattitudini costitutive dell'identità – la rimozione del desiderio non è una soluzione accettabile. Tizia potrebbe trovarsi nella situazione descritta senza credere in alcun modo che vi siano delle ragioni normative (ragioni “esterne”) per cui lei, come ciascuna persona, debba fare quello che gli sembra giusto. Potrebbe pensare che le ragioni che riconosce sono il modo in cui fenomenologicamente si pongono le sue motivazioni: ragioni interne. Quando Tizia si chiede se un certo giudizio morale sia “giustificato” crede che l'unica risposta sensata riposi sulle sue attuali intuizioni morali sui fini. Questo vuol dire sottoscrivere il Metodo della Familiarizzazione.

Queste considerazioni rudimentali vogliono contribuire a gettare le basi per un'analisi dei rapporti tra ragionamento pratico e posizioni metaetiche. È importante tenere a mente – come si accennava in apertura – che, comunque, il Metodo della Familiarizzazione e il Metodo dell'Accettazione non pretendono di catturare le posizioni teoriche sul ragionamento pratico e sulla metaetica degli autori contemporanei analizzati. In primo luogo, molti autori potrebbero non avere adottato affatto un metodo esplicito⁶⁸. L'utilità di questa analisi consiste nel radicalizzare provvisoriamente i loro argomenti per dissotterrare certe ambiguità (“se si deve accettare la Conclusione Ripugnante, questo vuol dire che il senso comune non è mai arbitro della correttezza?”). In secondo luogo, è facile immaginare che molti degli autori che criticano l'accettazione di conclusioni contro-intuitive, e che dunque sarebbero ascrivibili entro il Metodo della Familiarizzazione, adottino posizioni realiste o rigettino, su altre basi, il nichilismo. Da questo punto di vista, il dubbio che la presente analisi introduce è se effettivamente le loro posizioni formino un insieme coerente.

68 Prendiamo ad esempio l'argomento di Arrhenius contro l'Accettazione della Conclusione Ripugnante. È possibile che in altri suoi scritti Arrhenius abbia difeso conclusioni morali contro-intuitive.

6. IL PROBLEMA DEGLI ATTEGGIAMENTI DI SECONDO LIVELLO

Prima di concludere, può essere utile soffermarsi su un problema. La possibilità di distinguere tra il Metodo della Familiarizzazione e il Metodo dell'Accettazione è messa a rischio dall'esistenza di una certa categoria di stati mentali.

Fino a ora, per presentare i due stili morali, ho detto che la differenza consiste nel fatto che il Metodo dell'Accettazione ammette una mossa che non sembrerebbe essere ammessa dal Metodo della Familiarizzazione: l'accettazione di conclusioni contro-intuitive. Avere ragioni per accettare conclusioni contro-intuitive vuol dire ammettere che abbiamo ragioni per essere persone diverse da quelle che siamo - persone con un insieme di proattitudini diverse da quelle che ci capita di avere. Il realismo, tramite il Metodo dell'accettazione, ammetterebbe la sussistenza di ragioni per essere persone diverse da quelle che siamo, per contro il nichilismo, tramite il Metodo della familiarizzazione, le negherebbe. Questa rappresentazione della differenza tra i due stili è però frettolosa.

Accade che le persone abbiano degli atteggiamenti di secondo livello critici rispetto a certi loro atteggiamenti di primo livello⁶⁹. È un semplice dato di fatto che una persona che trova contro-intuitive certe azioni può desiderare essere una persona che non le troverebbe contro-intuitive.

Un esempio che abbiamo incontrato è quello di Greene che si rammarica di non riuscire ad avvertire, come molti altri, il valore intrinseco della creazione di nuove vite. Il giudizio "la creazione di nuove vite ha valore intrinseco" è un giudizio familiare per Greene? In un certo senso "no", perché il giudizio è contro-intuitivo, e questo è il motivo del rammarico di Greene; in altro senso "sì", perché, appunto, Greene già vorrebbe avere un giudizio del genere.

Gli atteggiamenti di secondo livello critici confondono le acque dell'analisi che ho presentato, perché, da un lato, sono emozioni, intuizioni, proattitudini - pezzi della nostra identità pratica - e, come tali, secondo il Metodo della Familiarizzazione, rendono giustificato il giudizio su cui vertono, ma, dall'altro, sono giudizi che favorevoli alla riforma di certi nostri giudizi, e dunque all'accettazione di conclusioni contro-intuitive.

Se gli atteggiamenti di secondo livello hanno valore anche nel Metodo della Familiarizzazione, allora non è vero, come prima si era detto, che questo stile morale sia imperniato sul rigetto delle conclusioni moralmente contro-intuitive: anche per il Metodo della Familiarizzazione sarà sensato accettare una conclusione contro-intuitiva nella misura in cui abbiamo un atteggiamento di secondo livello in tal senso.

Mi sembra che comunque il collasso del Metodo della Familiarizzazione nel Metodo dell'Accettazione può essere evitato. Una cosa è dire che abbiamo delle ragioni per accettare conclusioni morali contro-intuitive, altra cosa che abbiamo

69 Frankfurt 1971.

delle ragioni per accettare delle conclusioni morali contro-intuitive *finché* abbiamo un'inclinazione in tal senso.

Ma non è forse vero che ogni proposta di accettare conclusioni morali controintuitive è sottoposta alla condizione implicita che l'accettazione rispecchi una qualche nostra inclinazione morale in tal senso? Ossia, l'accettazione delle conclusioni controintuitive non deve forse essere in qualche modo moralmente intuitiva? Mi sembra che questo sia il nodo che va sciolto da parte di coloro che hanno sdoganato l'accettazione delle conclusioni contro-intuitive. Questi autori dovrebbero chiarire quando e perché si devono accettare conclusioni contro-intuitive, al fine di evitare una rappresentazione caricaturale della loro posizione.

Se chi sostiene l'accettazione delle conclusioni controintuitive subordina l'accettazione alla sussistenza di una inclinazione in tal senso, allora non abbiamo prova che il ragionamento morale miri a trascendere le contingenze della psicologia umana: se alcune inclinazioni morali sono pregiudizi mentre altre non lo sono è una questione che viene da ultimo decisa sulla base delle nostre stesse, contingenti, inclinazioni morali. Se chi sostiene l'accettazione delle conclusioni controintuitive non subordina l'accettazione alle inclinazioni contingenti, allora abbiamo effettivamente due metodi del ragionamento morale diversi. Per il Metodo della Familiarizzazione le conclusioni morali contro-intuitive vanno accettate nella misura in cui abbiamo un'inclinazione morale in tal senso; per il Metodo dell'Accettazione può darsi il caso che dobbiamo accettare delle conclusioni morali contro-intuitive anche quando non abbiamo un'intuizione morale favorevole in tal senso (ma abbiamo intuizioni teoriche favorevoli di diverso genere).

Potrebbe obiettarsi che se questa è la differenza tra il Metodo della Familiarizzazione e il Metodo dell'Accettazione allora non siamo davanti a due metodi diversi del ragionamento morale comune. In fondo, normalmente quando le persone impiegano frasi del tipo “dobbiamo sforzarci di capire il loro dolore” o “dobbiamo sforzarci di non lasciarci commuovere” - ossia manifestano la volontà di accettare conclusioni contro-intuitive - è facile ipotizzare che lo facciano perché hanno un'inclinazione morale in tal senso. Mi sembra corretto. Il Metodo della Familiarizzazione sarebbe allora il modo di ragionare comune e il Metodo dell'Accettazione sarebbe tendenzialmente appannaggio dell'analisi filosofica. Il seguace del Metodo della Familiarizzazione sarebbe, dunque, in disaccordo solo con l'uso filosofico dell'accettazione delle conclusioni contro-intuitive e non anche con l'uso di senso comune. Una posizione che ricorda quella di Wittgenstein - nella lettura di Kripke -, che era contro l'idea di “seguire una regola” costruita dai filosofi, ma non anche contro la pratica comune del pensare di stare seguendo una regola.

7. CONCLUSIONI

Con questo articolo ho offerto una presentazione della crisi della morale, servendomi dell'opera recente di Christopher Cowie, che ha avuto il merito di convogliare e sviluppare alcune tensioni presenti nel dibattito morale contemporaneo. Oltre a riprendere l'itinerario di Cowie, ho proposto un modo alternativo di sviluppare questi problemi. In entrambi i casi – nella mia operazione e in quella di Cowie – il dibattito morale sostanziale viene collegato in modo stretto al dibattito metaetico. Per Cowie i problemi che affiorano nell'etica sostanziale hanno implicazioni, trascinano, nell'ambito metaetico, e spingono verso l'adozione della teoria dell'errore. L'itinerario prefigurato da me è inverso: certe posizioni metaetiche hanno implicazioni che riguardano l'etica sostanziale. L'argomento di Cowie è una critica – secondo lo stesso autore non decisiva – a certe posizioni metaetiche, mentre il mio si risolve in una critica a certi argomenti adoperati nell'ambito dell'etica sostanziale. In particolare, l'accettazione di conclusioni contro-intuitive potrebbe richiedere – quello presente è solo uno studio preliminare – l'adozione di una posizione metaetica realista. La critica non è che il realismo morale sia indifendibile. La critica di fondo è che il dibattito su questo nesso è sottosviluppato, e che è necessario svilupparlo, affinché nel dibattito morale si possa giocare “a carte scoperte”.

Non so se sia il mio percorso o quello di Cowie a essere più convincente, né se i due percorsi siano irriducibilmente alternativi. In ogni caso mi sembra che il problema dei paradossi spinga lo studio sulla normatività verso una maggiore integrazione tra metaetica ed etica sostanziale, da un lato, e, dall'altro, tra queste e la teoria del diritto, perché diventa cruciale capire – come osservato da Celano – cosa fare se la morale è in crisi; capire se le istituzioni giuridiche possano rappresentare una soluzione.