

# ETICA DELLA POPOLAZIONE PARADOSSI, INTUZIONI E METODO

GIUSEPPE **ROCCHÉ**



Etica della popolazione  
Paradossi, intuizioni e metodo

Population Ethics  
Paradoxes, intuitions, and method

GIUSEPPE ROCCHÉ

Dottore di ricerca, borsista.  
E-mail: [giusepperoch@hotmail.it](mailto:giusepperoch@hotmail.it)

#### ABSTRACT

L'articolo introduce alcuni dei problemi da cui ha avuto origine il dibattito contemporaneo di etica della popolazione. Inizialmente viene mostrato come, lungi dal riguardare soltanto l'Utilitarismo, questi problemi coinvolgono alcune intuizioni comuni difficili da abbandonare. Di seguito vengono analizzati diversi tentativi di soluzione che si mostrano inefficaci. In appendice vengono svolte alcune considerazioni critiche sul metodo del ragionamento morale. In proposito vengono delineati due metodi alternativi sulla base del se sia ammesso o meno accettare conclusioni contro-intuitive.

This paper is an introduction to the contemporary debate on population ethics. As it is shown at the beginning, the paradoxes of population ethics do not pose a problem just for Utilitarianism, but also for everyone who shares a set of very common intuitions. Different proposals for a solution are described, each of which proves ineffective. At the end of the paper two different methods of moral reasoning are distinguished depending on whether accepting counter-intuitive conclusions is allowed or not.

#### KEYWORDS

etica della popolazione, Derek Parfit, conclusione ripugnante, utilitarismo, intuizioni morali  
population ethics, Derek Parfit, repugnant conclusion, utilitarianism, moral intuitions

# Etica della popolazione

## Paradossi, intuizioni e metodo

GIUSEPPE ROCCHÉ

1. *Scelte determinanti il numero* – 2. *Il Paradosso della Mera Addizione* – 3. *Il Principio del Valore Variabile* – 4. *Il Principio del Livello Critico* – 5. *Il Principio Lessicale* – 6. *“Person-affecting moralities”. La restrizione personalista* – 7. *Intransitività, impossibilità* – 8. *Accettare la Conclusione Ripugnante* – 9. *Note su alcuni problemi di metodo nel ragionamento morale.*

### 1. *Scelte determinanti il numero*

Questo articolo si propone di offrire una panoramica del dibattito di etica della popolazione come sviluppatosi a partire dall’opera di Derk Parfit. In appendice verranno svolte alcune considerazioni critiche.

Le nostre scelte possono influire sul benessere delle persone che esisteranno nel futuro. Oltre a questo, le nostre scelte possono influire sul numero di persone che esisteranno nel futuro. Chiamiamo questo secondo tipo di scelte “*scelte determinanti il numero*”<sup>1</sup>. Quali principi governano le scelte determinanti il numero? L’etica della popolazione si occupa di fornire risposte a questo interrogativo.

Immaginiamo di dovere scegliere tra due popolazioni alternative

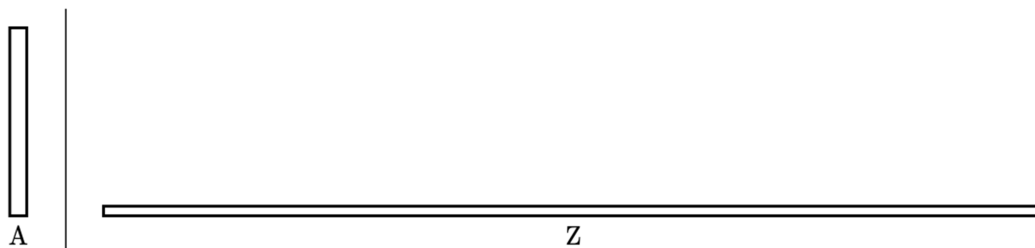


Fig.1

La base dei poligoni rappresenta il numero di persone esistenti, l’altezza rappresenta il livello di benessere goduto. A è una popolazione composta da esseri umani con un livello di benessere estremamente elevato. Per contro, Z è una popolazione enormemente più grande – possiamo immaginare una popolazione anche ben maggiore di come, per limiti di spazio, l’ho rappresentata graficamente –, in cui ciascuno ha una vita che, in termini di benessere<sup>2</sup>, può dirsi degna di essere vissuta, ma che è notevolmente peggiore rispetto a quelle di coloro che vivono in A: in Z le vite sono *appena* degne di essere vissute – più prosaicamente, in Z ciascuno vive una vita in cui la somma algebrica delle componenti positive e negative del benessere ha un risultato che è appena positivo<sup>3</sup>. A non rappresenta una popolazione esigua per i nostri standard. Possiamo

<sup>1</sup> PARFIT 1984, 355 s.

<sup>2</sup> Non possiamo soffermarci sull’estensione del concetto di benessere, per quanto pertinente sia il tema. Sulle varie concezioni del benessere un classico è rappresentato da GRIFFIN 1986, 8-72.

<sup>3</sup> Se pensiamo a migliaia di miliardi di persone che vivono contemporaneamente sulla Terra e non abbiamo un’immaginazione fantascientifica particolarmente raffinata, potremmo ritrovarci a immaginare un inferno. Que-

immaginarla come una popolazione di svariati miliardi di esseri umani. Sia in A sia in Z regna l'eguaglianza perfetta.

La domanda che anima il dibattito di etica della popolazione è: Sarebbe meglio che esistesse A o Z? Ossia: ne avessimo le possibilità, dovremmo produrre A o Z?

Molte persone tendono a dire che A sia preferibile a Z.

Verosimilmente A non è sempre preferibile a Z. Immaginiamo che A sia un mondo in cui esistono solo persone malvagie, mentre Z sia composto da persone oneste e altruiste. Se tali fossero A e Z, potrebbero esservi ragioni per giudicare Z migliore di A. L'esempio illustra come il benessere non è l'unica variabile da tenere in considerazione nell'elaborare giudizi pratici. Ma il giudizio che qui interessa è che A sia preferibile a Z quando le altre variabili che potrebbero influire sulla nostra valutazione sono ritenute irrilevanti – clausola *ceteris paribus*.

L'impiego della clausola *ceteris paribus* non è esente da complicazioni. Si dice, ad esempio, che le due popolazioni alternative non fanno differenza per quanto concerne il rispetto dei diritti fondamentali<sup>4</sup>. Ebbene, è difficile credere che A possa essere un mondo in cui il benessere è altissimo ma i diritti umani sono calpestati. Il rispetto dei diritti fondamentali potrebbe essere una condizione per la realizzazione di una società gioiosa come A pretende di essere. Se in A i diritti fondamentali sono pienamente rispettati, per la clausola *ceteris paribus* devono esserlo anche in Z. Z sarebbe allora un mondo in cui il benessere è molto basso ma i diritti fondamentali sono perfettamente realizzati. Ma è difficile immaginare un mondo del genere per la ragione speculare a quella che ci impediva di immaginare A come un mondo gioioso in cui i diritti fondamentali sono calpestati. Se in Z le libertà civili, i poteri politici e la sicurezza sociale sono garantiti a ciascuno quanto in A, com'è possibile che il benessere sia così basso?<sup>5</sup> La clausola *ceteris paribus* potrebbe non essere facilmente gestibile. Data la finalità introduttiva di questo articolo, non si insisterà ulteriormente su questo punto.

Secondo l'Utilitarismo Classico o Principio di Utilità Totale (PUT), dati due o più mondi possibili, il mondo migliore è quello che contiene la quantità maggiore di benessere<sup>6</sup>. Sebbene in

sta rappresentazione tradirebbe il senso di Z. Stiamo, infatti, adottando come premessa che ciascuno in Z vive una vita che è appena degna di essere vissuta. Se abbiamo difficoltà a immaginare migliaia di miliardi di persone che vivono nello stesso pianeta e che hanno vite comunque – sebbene appena – degne di essere vissute, possiamo immaginare Z come un insieme di migliaia di pianeti, ciascuno popolato da miliardi di persone, in cui ciascuno vive una vita appena degna di essere vissuta.

<sup>4</sup> ARRHENIUS 2000b, 42.

<sup>5</sup> In Z ciascuno va in scuole identiche a quelle presenti in A – identica deve essere la realizzazione del diritto all'istruzione – e le scuole pubbliche di A sono scuole perfette in cui a ciascuno viene insegnato a vivere bene, dunque anche chi vive in Z va in scuole del genere e gli viene insegnato a vivere bene; in Z ciascuno godrà di un sistema previdenziale massimamente rassicurante come quello presente in A; in Z ciascuno sarà messo nelle condizioni di – ed attivamente spronato a – sviluppare la propria coscienza critica, le proprie attitudini pro-sociali (studiare, accedere ad attività culturali, ecc.) nello stesso modo in cui ciò avviene in A; e così via. Salvo non si adotti una concezione minimalista dei diritti fondamentali (contro il minimalismo CELANO 2013, 85 ss.), sembra esservi un rapporto estremamente stretto tra rispetto, pieno, dei diritti fondamentali e benessere individuale: per cui (salvo forse casi eccezionali) non possono darsi due mondi alternativi, uno in cui il benessere è molto alto e l'altro in cui il benessere è molto basso, in cui i diritti fondamentali sono garantiti allo stesso modo – ossia, sono compiutamente realizzati in entrambi oppure sono totalmente calpestati in entrambi.

<sup>6</sup> Seguendo l'uso di Parfit, l'Utilitarismo Classico adotterebbe il Principio Totale come principio aggregativo del benessere (PARFIT 1984, 449). In ciò si distinguerebbe da un'altra versione dell'Utilitarismo che utilizzerebbe il Principio della Media. Secondo quest'ultima versione, non sarebbe giusto massimizzare il benessere totale, ma massimizzare il benessere medio (pro capite) della popolazione (sulla distinzione tra Principio Totale e Principio della Media, PARFIT 1984, 384 ss.; vi torneremo subito §2). Parfit non è solo nell'identificare l'Utilitarismo Classico con il Principio Totale, ma segue una terminologia che si ritrova in vari autori (così SIKORA 1975, 409 s.), tra cui RAWLS 1971, 161 ss. Jan Narveson, d'altra parte, ritiene che sia frutto di una confusione credere che l'Utilitarismo Classico imponga agli agenti morali di aumentare il numero delle persone con benessere positivo, ad esempio facendo più figli possibili (NARVESON 1967, 62 ss.). Secondo Narveson «when we [gli utilitaristi] specify the individuals who would be affected by our actions, as we must on the utilitarian view, the characteristic about those people with

Z ciascuno sta peggio di chiunque viva in A, il numero di persone che vive in Z è talmente grande da contenere una quantità di benessere maggiore di quella di A. Un mare di bicchieri contenenti ciascuno una sola goccia d'acqua può contenere molta più acqua di una cisterna piena fino all'orlo. Per PUT, dunque, Z è meglio di A. PUT implica allora

la *Conclusionone Ripugnante*. Per ogni popolazione possibile composta da miliardi di persone, ciascuna delle quali ha un livello di benessere molto elevato, è possibile immaginare una popolazione molto maggiore la cui esistenza, a parità di condizioni, sarebbe migliore anche se ciascuno dei suoi membri conduce una vita che è appena degna di essere vissuta<sup>7</sup>.

L'Utilitarismo è un'ideologia morale estremamente esigente che onera l'agente di mettersi al servizio del bene comune fino all'estremo sacrificio<sup>8</sup>. Di più, ciascuno, in quanto operatore al servizio del bene comune, avrebbe il dovere morale di sacrificare alla causa universale non solo il proprio benessere ma anche quello altrui<sup>9</sup>. Queste sono obiezioni ben note. Che l'Utilitarismo Classico implichi che Z sia preferibile ad A può essere vista come un'ulteriore obiezione nei confronti del medesimo. Operando le dovute semplificazioni, immaginiamo che un modo per incrementare il benessere contenuto in una vita sia dato dal prolungare quella vita. Possiamo, allora, immaginare che la differenza in termini di benessere tra A e Z sia data dal fatto che l'aspettativa di vita media in A è decisamente più lunga di quella in Z. Abbiamo due alternative.

Prima politica: *Curare*:  $x$  risorse pubbliche verranno destinate al sistema sanitario nazionale. In conseguenza di ciò molte persone che svilupperanno mali potenzialmente letali o gravemente debilitanti, ma facilmente curabili – ad esempio fratture ossee, appendiciti, ecc. –, verranno efficacemente trattate – mondo A.

Seconda politica: *Rimpiazzare*: le  $x$  risorse che sarebbero state destinate al sistema sanitario nazionale vengono invece destinate a incentivare le nascite. A seguito di ciò molte persone muoiono giovani o patiscono gravi debilitazioni per via di mali facilmente curabili. Nel complesso, però, esistono tantissime persone in più – mondo Z<sup>10</sup>.

which we are morally concerned is whether their happiness will be increased or decreased [...] The child cannot be happier as a result of being born, since we would then have a relative term lacking one relatum. The child's happiness has not been increased, in any intelligible sense, as a result of his being born; and since nobody else's has either, directly, there is no moral reason for bringing him into existence» (NARVESON 1967, 65, 67). Chiamare – seguendo Rawls e Parfit, ma non Narveson – il Principio Totale “Utilitarismo Classico” corrisponde a un uso recente, ma questo uso non sottende una tesi storica sulla posizione dei primi grandi utilitaristi. Ad esempio, Rawls (1971, 162) afferma che J.S. Mill sia stato un sostenitore del Principio della Media – Mill, dunque, non sarebbe un utilitarista classico in questo senso (e almeno sul fatto che non sia, in questo senso, un utilitarista classico Rawls e Narveson sarebbero d'accordo). Questioni terminologiche a parte, si potrebbe pensare, seguendo Narveson, che un'idea forte di ispirazione utilitarista sia quella per cui bisogna massimizzare il benessere delle persone che esistono. Questa idea narvesoniana – che è stata prospettata da un *referee* – merita certamente discussione e verrà analizzata di seguito (v. §6).

<sup>7</sup> PARFIT 1984, 388. Parfit in realtà parla non di “benessere” ma di “qualità della vita”. A questa espressione viene da lui attribuito un significato specifico che vedremo in §5 quando si analizzerà il valore del Perfezionismo.

<sup>8</sup> Per questa critica WILLIAMS 1973b; ma anche SCHEFFLER 1982. Un ampliamento della sfera dei nostri doveri morali è stato invece difeso in SINGER 1972, e UNGER 1996.

<sup>9</sup> Così, un chirurgo avrebbe il dovere di circuire e fare a pezzi un ignaro passante, qualora gli organi così ottenuti gli permettessero di salvare la vita a cinque moribondi. Un classico del dibattito su questa posizione THOMSON 1985, recentemente GREENE 2013.

<sup>10</sup> Insomma l'alternativa è tra creare nuove persone e salvare la vita alle persone già esistenti, così MCMAHAN 2013, appunto «*Causing People to Exist and Saving People's Lives*», in particolare p. 31 s.

L'Utilitarismo Classico giustifica l'adozione della seconda politica. Un esito contro-intuitivo che si coglie una volta che consideriamo le scelte determinanti il numero. Si tratta di un'obiezione indipendente, aggiuntiva rispetto alle precedenti, ma anche un'obiezione che in qualche modo le rafforza: chiedere al singolo di essere pronto a sacrificare se stesso e il proprio prossimo può essere già eccessivo quando il fine è rappresentato dalla tutela del benessere di altre persone che esistono da qualche parte sul pianeta e che altrimenti patirebbero gravi danni, ma lo è ancora di più quando è rappresentato dalla creazione di nuove persone che altrimenti non sarebbero esistite.

## 2. Il Paradosso della Mera Addizione

Come abbiamo appena visto, l'Utilitarismo Classico, PUT, implica quello che Parfit ha chiamato "Conclusione Ripugnante". Nel passato alcuni autori avevano notato questa deriva. Più recentemente simili problemi sono stati affrontati da Jan Narveson, ritenuto il pioniere dell'etica della popolazione contemporanea<sup>11</sup>. Tuttavia, è certamente con il lavoro di Parfit dei primi anni 80 che l'etica della popolazione è emersa come un nuovo ambito del dibattito morale<sup>12</sup>.

In proposito, una delle più originali e feconde mosse di Parfit è consistita nel mostrare come la Conclusione Ripugnante non costituisca un problema solo per i sostenitori dell'Utilitarismo Classico (PUT): sostenere che A sia meglio di Z si rivela più difficile del previsto<sup>13</sup>. Consideriamo le seguenti popolazioni alternative rappresentate nella fig. 2

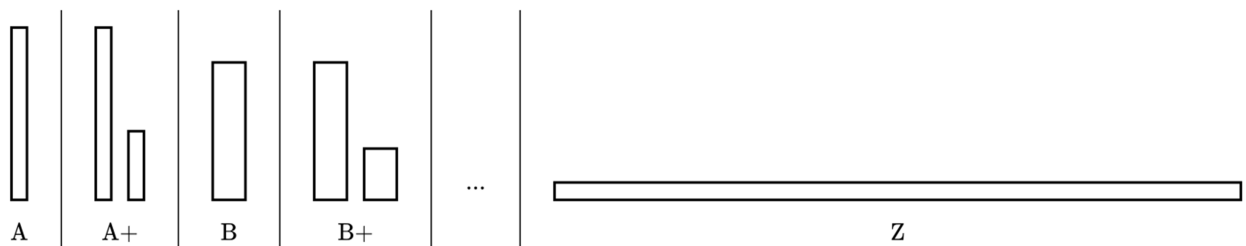


Fig. 2

A è una popolazione in cui esistono miliardi di persone con un livello di benessere molto elevato. In A+ altrettante persone<sup>14</sup> vivono ad un livello di benessere uguale a quello di A. Però, in A+ esistono anche molte altre persone aggiuntive con un livello di benessere inferiore, ma comun-

<sup>11</sup> Sidgwick nota come l'Utilitarismo non prescriva – come pensavano i malthusiani – di incoraggiare gli aumenti della popolazione solo fino al punto in cui il benessere medio è il massimo possibile, ma, viceversa, ammetta che la popolazione possa crescere ancora di più anche se ciò determinerebbe un decremento del benessere medio (SIDGWICK, 1907, 415). McTaggart accetta una versione della Conclusione Ripugnante, rilevando come la ripugnanza nel caso di specie sia cattiva consigliera (MCTAGGART 1921, 452 s.). Per l'opera di Jan Narveson si vedano NARVESON 1967 e NARVESON 1973.

<sup>12</sup> Come scrive Jeff McMahan nel necrologio dedicato al suo caro amico Parfit «While each of the four parts [di *Reasons and Persons*, la più celebre opera di Parfit] has been enormously influential. Part Four effectively created a new, difficult, highly important, and now flourishing area of philosophy known as “population ethics”». In Italia questi temi hanno ricevuto limitata attenzione, un'eccezione è rappresentata da PELLEGRINO 2012.

<sup>13</sup> «Suppose that we do reject this principle [il Principio di Utilità Totale]. Unfortunately, this is not enough» PARFIT 1986, 151.

<sup>14</sup> A volte, nella presentazione di A+ gli autori pongono il caso che A+ contenga le *stesse* persone che esistono in A (con il medesimo livello di benessere), più un certo numero di persone con un livello di benessere inferiore. Questo dettaglio potrebbe fare la differenza in base ai principi che verranno analizzati in §6. Dicendo che A+ contiene *altrettante* persone di quante sono presenti in A, ho scelto di non inserire questo dettaglio per non complicare l'introduzione al problema, seguendo in questo la formulazione scelta da Parfit (PARFIT 1984, 419 s.).

que positivo. La popolazione  $A+$  è frutto di una *mera addizione* rispetto ad  $A$ , ossia in  $A+$  è presente un gruppo di persone aggiuntive (“extra group”) (i) che hanno vite comunque degne di essere vissute, (ii) la cui aggiunta non determina un danno per nessuno, (iii) e che non sono vittime di ingiustizia sociale<sup>15</sup>.

Che  $Z$  sia migliore di  $A$  è la Conclusione Ripugnante. Diciamo che il *Rigetto della Conclusione Ripugnante* costituisce una prima condizione di accettabilità di una teoria o principio morale.

$B$ , a sua volta, sembra essere migliore di  $A+$ . In primo luogo, uno degli aspetti negativi lampanti di  $A+$  è costituito dalla presenza di accentuate diseguaglianze nei livelli di benessere. In  $B$ , invece regna l’eguaglianza perfetta. Questa considerazione potrebbe non essere decisiva. L’eguaglianza perfetta tra persone del tutto sane e persone zoppe può essere realizzata azzoppando le persone sane. Molti ritengono che ciò non costituirebbe un miglioramento. Tuttavia,  $B$  non è il prodotto del livellamento verso il basso: coloro che vivono in  $B$  stanno meglio dei più sfortunati di  $A+$ .

$B$  è preferibile rispetto al principio del Maximin, per cui la popolazione migliore è quella in cui i più svantaggiati stanno meglio. Non sempre, potrebbe dirsi, la redistribuzione del benessere a favore dei più svantaggiati corrisponde ad un miglioramento dello scenario considerato. Immaginiamo una società molto opulenta in cui vi è un piccolo gruppo notevolmente meno abbiente degli altri, che, però, vive in condizioni già ampiamente accettabili. Se immaginiamo che l’unico modo per migliorare *appena* le condizioni di vita dei meno abbienti sia diminuire enormemente le condizioni di vita di tutti gli altri membri della società, alcuni potrebbero ritenere che preservare lo *status quo* sia preferibile. Tuttavia, nemmeno questo è il rapporto tra  $A+$  e  $B$ <sup>16</sup>. Significativamente, il benessere medio e totale in  $A+$  è inferiore a quello in  $B$ : i più svantaggiati guadagnano più di quanto i più avvantaggiati perdono.

Secondo il *Principio di Non-Antieguaglianza* una popolazione perfettamente egualitaria è meglio di una pari popolazione che presenta diseguaglianze e che ha un livello medio (e dunque totale) di benessere inferiore<sup>17</sup>. Il Principio di Non-Antieguaglianza si limita a sostenere che l’innalzamento del livello di benessere dei più sfortunati che non si risolve in un pregiudizio per il benessere totale dell’intera comunità ha un valore positivo. Si tratta di un principio chiaramente attraente. Il fatto che una teoria morale, o un principio, non abbia esiti anti-egualitari – che rispetti il Principio di Non-Antieguaglianza – sembra essere una condizione di ammissibilità di quella teoria o principio.  $B$  sembra dunque migliore di  $A+$ .

$A+$ , tuttavia, appare *almeno tanto buono quanto*  $A$ <sup>18</sup>.  $A+$  è peggiore di  $A$  secondo il Principio di Utilità Media, per cui la popolazione migliore è quella con il benessere medio (pro capite) più elevato. Il Principio della Media, tuttavia, dà esiti contro-intuitivi. Immaginiamo che una coppia si stia interrogando sull’opportunità di avere un figlio. Seguendo il Principio della Media, la coppia dovrebbe chiedersi non solo quale sarà verosimilmente il livello di benessere del figlio, ma anche la variazione del benessere medio che l’aggiunzione di questa nuova persona determinerebbe. Così, vi sarebbe una ragione per astenersi dall’aver un figlio che avrà un livello di benessere molto elevato, qualora il benessere medio delle *altre* persone sia ancora più elevato; e, specularmente, potrebbe esservi una ragione per avere un figlio che avrà un livello di benessere negativo – una vita tormentata – qualora il benessere medio delle altre persone sia ancora più basso<sup>19</sup> – variabilità del valore contributivo di una vita, ci torneremo in §3.

<sup>15</sup> Potrebbe aversi ingiustizia sociale qualora la diseguaglianza fosse nota al gruppo avvantaggiato e rimediabile. Parfit non si dilunga sul punto (PARFIT 1984, 419 s.).

<sup>16</sup> PARFIT 1984, 422, 425 ss. Il Maximin nella filosofia politico-morale della seconda metà del ‘900 svolge un ruolo chiave nella teoria della giustizia di Rawls (RAWLS 1971).

<sup>17</sup> Così in NG 1989, 238. La formulazione qua usata è tratta da ARRHENIUS 2000b, 70.

<sup>18</sup> Parfit in proposito usa l’espressione “*non peggiore*”, con l’intento di sollevare un problema di vaghezza nella comparazione tra  $A$  e  $A+$  (PARFIT 1984, 426, 430-432). Nell’usare l’espressione “*almeno tanto buono quanto*” nel contesto della presentazione del problema, mi allineo ad ARRHENIUS 2016, 175.

<sup>19</sup> PARFIT 1984, 420-422.

A+ potrebbe essere peggiore rispetto al principio del Maximin, per cui la popolazione migliore è quella in cui i più svantaggiati stanno meglio – si noti che in A, regnando l'eguaglianza, i più svantaggiati corrispondono ai più avvantaggiati. Questa conclusione, tuttavia, non sembra solida. Se prendiamo in esame due scenari con lo stesso numero di persone, lo scenario in cui i più svantaggiati stanno meglio è anche lo scenario in cui vi sono *più persone* che stanno meglio rispetto ai più svantaggiati dell'altro scenario. Se, in un ipotetico scenario X, metà della popolazione ha benessere 100 e l'altra metà ha benessere 50, mentre nello scenario Y metà della popolazione ha benessere 100 e l'altra metà 20, allora è vero sia che nella prima popolazione i più svantaggiati stanno meglio, sia che nella prima popolazione vi sono più persone che stanno meglio di chi ha benessere 20 nel secondo scenario. Ciò non è però vero quando consideriamo casi di mera addizione. Sebbene sia vero che in A tutti stanno meglio rispetto ai più svantaggiati in A+, non è vero che in A vi sono più persone che stanno meglio dei più svantaggiati in A+; invero, l'esistenza dei più svantaggiati in A+ è semplicemente frutto del fatto che la popolazione A+ è maggiore di quella di A.

Ora, che in uno scenario vi siano più persone che stanno meglio dei più svantaggiati dell'altro scenario è chiaramente una buona caratteristica di quello scenario. In questi casi, comparazione X *versus* Y, il Maximin ha rilevanza – possiamo dire anche che il Maximin riconosce salienza alla circostanza che in Y c'è chi ha un livello di benessere inferiore a quello di chi vive in X. Ma che in uno scenario il benessere dei più svantaggiati sia più elevato che nell'altro, solo perché si è fatto a meno di creare altre persone possibili che avrebbero avuto una vita degna di essere vissuta, non sembra una buona caratteristica di quello scenario. In questi casi, comparazione A *versus* A+, il Maximin non ha rilevanza<sup>20</sup> – oppure il Maximin *non* riconosce salienza alla circostanza che in A+ c'è chi ha un livello di benessere inferiore a quello di chi vive in A<sup>21</sup>.

Chiamiamo “*Principio della Mera Addizione*” il principio secondo il quale la semplice aggiunta di persone con vite degne di essere vissute non può peggiorare il valore morale di uno scenario<sup>22</sup> – A+ è almeno tanto buono quanto A. Pure il Principio della Mera Addizione sembrerà a molti essere una condizione di accettabilità di ogni teoria o principio morale degno di considerazione.

Se B è migliore di A+, e A+ è almeno tanto buono quanto A, segue che B è anche migliore di A. Possiamo ripetere il processo che abbiamo visto con la popolazione B: effettuando una mera addizione, otteniamo una popolazione B+ che sarà almeno tanto buona quanto B; vi sarà poi una popolazione C che sarà migliore di B+ e dunque migliore di B e di A. Possiamo ripetere il processo con C e poi con D e così via fino a trovare una popolazione Z che sarà migliore di A<sup>23</sup>. Otterremmo allora che Z sia migliore di A, ma questo equivale a violare il requisito di accettabilità che ci eravamo posti in partenza, ossia il *Rigetto della Conclusione Ripugnante*.

I passaggi intermedi tra A, A+ e Z sono stati riportati per descrivere il problema a cui stiamo andando incontro quasi nei termini originari con cui venne posto da Parfit<sup>24</sup>. A scopo semplificato, possiamo ignorare i passaggi intermedi, e fare riferimento d'ora in avanti alla seguente comparazione (Fig.3), dove, per le ragioni esposte precedentemente, A+ sarà almeno tanto buono quanto A e Z (non più B) sarà migliore di A+:

<sup>20</sup> PARFIT 1984, 422 ss.

<sup>21</sup> Abbiamo due concezioni rivali sul punto. Secondo la prima il Maximin non avrebbe rilevanza nella comparazione tra A e A+ – questa è la concezione che sembra adottare Parfit. In base alla seconda concezione il Maximin rilevarebbe sempre ma non riconoscerebbe salienza alla circostanza che in A+ ci sono persone con un livello di benessere inferiore a quello delle persone che vivono in A – questa concezione si ritrova in TEMKIN 2012, 404. Peraltro Temkin argomenta come la distinzione non abbia grande rilevanza pratica, TEMKIN 2012, 445 ss.

<sup>22</sup> NG 1989, 238; ARRHENIUS 2000b, 60.

<sup>23</sup> ARRHENIUS 2016, 175.

<sup>24</sup> Quella riportata è “quasi” la versione originale del problema. Infatti, secondo Parfit A+ e A sono solo vagamente l'uno non peggiore dell'altro (v. nt.18).



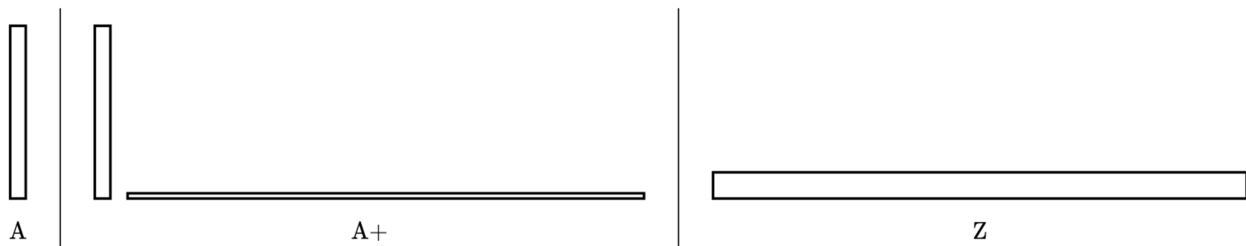


Fig. 3

Abbiamo dunque tre requisiti di accettabilità di una plausibile teoria morale: (i) il Rigtetto della Conclusione Ripugnante; (ii) il Principio di Non-Antieguagliarismo; (iii) il Principio della Mera Addizione. Sebbene ciascuno di questi tre principi abbia grande plausibilità, insieme considerati si rivelano contraddittori. Infatti, per via del Rigtetto della Conclusione Ripugnante, A è meglio di Z; per via del Principio di Non-Antieguagliarismo, Z è meglio di A+; ma, per il Principio della Mera Addizione, A+ è almeno tanto buono quanto A. Se teniamo fermo il Rigtetto della Conclusione Ripugnante – A è meglio di Z – e il Principio di Non-Antieguagliarismo – Z è meglio di A+ –, allora, per *transitività*, dovrebbe seguire che A+ sia peggiore di A – ma questo equivarrebbe a negare il Principio della Mera Addizione. D'altra parte, se teniamo fermo il Principio della Mera Addizione – A+ è almeno tanto buono quanto A – allora, per *transitività* o *sostituzione*, qualsiasi scenario migliore di A+ sarà migliore anche di A, e qualsiasi scenario peggiore di A sarà peggiore anche di A+. Mettiamo, allora, di voler preservare il Principio di Non-Antieguagliarismo: essendo Z migliore di A+, seguirebbe che Z è migliore di A – avremmo allora violato la condizione del Rigtetto della Conclusione Ripugnante. Mettiamo, invece, di voler preservare il Rigtetto della Conclusione Ripugnante: essendo Z peggiore di A, seguirebbe anche che Z sia peggiore di A+ – ma ciò equivale alla violazione del Principio di Non-Antieguagliarismo<sup>25</sup>.

Conclusione. Chiamiamo *Paradosso della Mera Addizione* la dimostrazione della contraddittorietà di queste tre intuizioni comuni – (i) il Rigtetto della Conclusione Ripugnante; (ii) il Principio di Non-Antieguagliarismo; (iii) il Principio della Mera Addizione<sup>26</sup>. Una di queste intuizioni deve essere abbandonata. Ma quale?

Una precisazione. Il Paradosso della Mera Addizione parrebbe essere un *conflitto* tra valori. Invece, secondo Parfit, esso illustra una *contraddizione* tra giudizi morali. Chiamiamo. Quando vi è una pluralità di principi o valori in gioco è possibile che questi confliggano: l'azione X è sorretta dal principio P<sub>1</sub>, ma va contro il principio P<sub>2</sub>. Quando si manifestano conflitti di questo genere possiamo trovarci spaesati – non sapere se dobbiamo o no fare X – senza che i nostri giudizi siano contraddittori. Per quanto difficile possa sembrare in certi casi, dobbiamo chiarirci se X, che è *prima facie* una buona azione per via di P<sub>1</sub>, sia anche una buona azione *tutto considerato* – chiarirci se il danno in termini di P<sub>2</sub> sia più che compensato dal vantaggio in termini di P<sub>1</sub>. Nel Paradosso della Mera Addizione le cose non stanno così. Quando, ad esempio, diciamo che, per il Principio della Mera Addizione, A+ è almeno tanto buono quanto A, non vogliamo dire che *prima facie* A+ è almeno tanto buono quanto A rispetto al valore del benessere totale, ma potrebbe non esserlo *tutto considerato* perché A è meglio di A+ rispetto, ad

<sup>25</sup> Seguendo TEMKIN (1987, 143 s.), la relazione “x è meglio di y secondo il valore c” è transitiva nel senso che «Given any concept c, for all x, y, and z to which c is appropriately applied, if x is c-better than y, and y is c-better than z, then x is c better than z». La relazione “x è almeno buono quanto y secondo il valore c” sussiste se alternativamente sussiste la relazione “x è meglio di y secondo il valore c” o “x è equivalente a y secondo il valore c”. La relazione “x è equivalente a y secondo il valore c” è sottoposta al *principio di sostituzione*, che recita «Given any concept c, for all x, y, and z to which c is appropriately applied, if x is as c-good as, or c-equivalent to y, then however x c compares to z, that is how y c-compares to z» (anche in TEMKIN 2012, 368). Dunque, la relazione “x è almeno buono quanto y secondo il valore c” permette di fare la seguente inferenza «If x is c-better than y, and y is at least as c-good as z, then x is c-better than z». Sulla *transitività* torneremo in §7.

<sup>26</sup> PARFIT 1984, 425 ss.

esempio, al Maximin. Quando diciamo che, per il Principio della Mera Addizione,  $A+$  è almeno tanto buono quanto  $A$ , vogliamo dire che, mentre vi sono considerazioni a favore di  $A+$  contro  $A$  – in  $A+$  il benessere totale è maggiore che in  $A$  –, non vi sono considerazioni a favore di  $A$  contro  $A+$  – come abbiamo visto sopra, il Maximin non è moralmente rilevante in presenza di una mera addizione. In altri termini,  $A+$  sembra essere almeno tanto buono quanto  $A$  tutto considerato. E così anche  $A$  sembra essere migliore tutto considerato rispetto a  $Z$ , e  $Z$  rispetto a  $A+$ <sup>27</sup>.

Secondo l'Utilitarismo Classico (PUT) possiamo fare del bene creando nuove vite con benessere positivo. PUT implica la Conclusione Ripugnante; ciò può essere inteso come una confutazione di PUT. Tuttavia, come abbiamo visto in questo paragrafo, il Paradosso della Mera Addizione non presuppone che vi sia valore nella creazione di vite con benessere positivo:  $A+$  è semplicemente almeno tanto buono quanto  $A$ . La lezione appresa è, pertanto, che per rigettare la Conclusione Ripugnante non possiamo solo prendere le distanze dall'Utilitarismo Classico – PUT –, ma sembrerebbe che dovremmo anche abbandonare o il Principio della Mera Addizione, o Non-Antiegalitarismo. I problemi dell'etica della popolazione non riguardano solo l'Utilitarismo<sup>28</sup>.

Dal prossimo paragrafo esamineremo alcuni tentativi di soluzione del problema.

<sup>27</sup> PARFIT 1984, 427; per una riconsiderazione del Paradosso della Mera Addizione in chiave di conflitto nella cornice di una morale pluralista TEMKIN 2012, 366, e 394 ss. Tuttavia forse una parziale abiura di questa distinzione si ritrova proprio in uno scritto tardo di Parfit in difesa del Perfezionismo (PARFIT 2016, 123 ss.), avremo modo di tornare sul punto §5.

<sup>28</sup> Certe persone potrebbero ritenere che  $A$  sia migliore di  $Z$ , ma che  $B$  (Fig.2) sia migliore di  $A$ : un mondo di miliardi di persone con benessere molto alto è meglio di un mondo composto da molte meno persone con un livello di benessere leggermente ancora più alto. Per queste persone la prima versione considerata del Paradosso della Mera Addizione – quella con la reiterazione delle mere addizioni (Fig.2, appunto) – nasconde un problema diverso da quello che si presenta nella seconda versione (Fig.3). Nella seconda versione, per loro, avremmo una tensione tra il Principio di Mera Addizione, il Principio di Non-Antiegalitarismo e il Rigo della Conclusione Ripugnante. Nella prima, invece, il Principio di Mera Addizione e il Principio di Non-Antiegalitarismo fungerebbero da argomenti in favore di una conclusione ben accetta in sé e per sé: che  $B$  sia migliore di  $A$ . Il problema consisterebbe allora nella reiterazione delle mere addizioni seguite da redistribuzioni del benessere, che condurrebbero, per transitività-sostituzione, a concludere che  $Z$  sia migliore di  $A$ . In questo caso, avremo un *Problema a Spettro*:  $B$  è migliore di  $A$ , e pure  $C$  è migliore di  $B$ ,  $D$  è poi migliore di  $C$ , e così via, però  $Z$  è peggiore di  $A$ . Qualcuno potrebbe credere che un simile Problema a Spettro sia analogo al *Paradosso del Sorite*. In una versione nota del Paradosso del Sorite, può mostrarsi come una persona sia allo stesso tempo capelluta e calva – esito paradossale. La premessa è che un singolo capello non può fare la differenza sul se una certa persona sia capelluta o calva. A questo punto si immagina uno spettro di mutazioni: a un estremo vi è una persona decisamente capelluta; segue una sequenza per cui, in ciascun stato, la persona perde un capello; all'altro estremo vi è una persona completamente calva. Ebbene, se rimuoviamo un capello dalla persona capelluta, è vero che quella persona resta capelluta. E lo stesso sarà vero se reiteriamo la rimozione di singoli capelli – infatti, la premessa da cui siamo partiti è che un singolo capello non può fare la differenza sul se una persona sia capelluta o meno. Così facendo, tuttavia, percorreremo tutto lo spettro delle mutazioni fino ad arrivare a uno stato in cui la persona è completamente calva, eppure la si dovrà dire ancora capelluta. Una premessa convincente – un singolo capello non può fare la differenza sul se una certa persona sia capelluta o calva – ha generato un esito assurdo – una persona è, allo stesso tempo, capelluta e calva. Il Problema a Spettro però non è affatto analogo al Paradosso del Sorite. Nel Paradosso del Sorite, per via della vaghezza insita nel concetto di “persona capelluta”, possiamo dire che una persona capelluta che perde un capello resta ancora una persona capelluta – sebbene patisca la perdita di un capello la persona capelluta resta una persona capelluta. Nel nostro problema a Spettro le cose sono completamente diverse. Per coloro che trovano  $B$  migliore di  $A$ , ma  $Z$  peggiore di  $A$ , man mano che si procede lungo lo spettro – aumentando di molto la popolazione e diminuendo leggermente il benessere medio –, il valore morale della popolazione va aumentando, però – questo l'esito paradossale – arrivati alla fine dello spettro ( $Z$ ) il valore morale è inferiore a quello della popolazione di partenza ( $A$ ). Sarebbe come se dicessimo che ogni qual volta rimuoviamo un capello alla persona capelluta questa diventa ancora più capelluta. Uno schema diverso da quello del Paradosso del Sorite – là, ogni volta che rimuoviamo un capello, la persona resta ancora una persona capelluta, ma non certo una persona più capelluta di prima – e certamente ben più misterioso (TEMKIN 2012, 277-284).

### 3. Il Principio del Valore Variabile

Chiamiamo “*valore contributivo di una vita*” il valore conferito a una data popolazione dall’aggiunzione di quella vita, e “*valore personale di una vita*” il risultato della somma algebrica delle componenti positive e negative del benessere<sup>29</sup>. Secondo PUT il valore contributivo di una vita è pari al valore personale – se ed, eventualmente, quanto sia positivo dipenderà dal risultato della somma algebrica – e non dipende dal livello di benessere delle altre persone che esistono. Chiamiamo questa caratteristica di PUT “*Invarianza del valore contributivo*”. Il Principio della Media si basa su una logica opposta: il valore contributivo varia a seconda del benessere delle altre persone che esistono; l’aggiunzione di una vita con benessere anche molto alto ha valore contributivo negativo qualora riduca il benessere medio della popolazione<sup>30</sup>. Anche le varie versioni del Principio del Valore Variabile negano – al pari del Principio della Media – l’Invarianza del Valore Contributivo: per stabilire il valore dell’aggiunzione di una persona dobbiamo guardare al benessere, del tutto indipendente, delle altre persone che compongono la popolazione di riferimento. Un tratto distintivo delle varie versioni del Principio del Valore Variabile risiede nell’attribuire, in modi diversi, *valore marginale decrescente* all’aggiunzione di vite<sup>31</sup> con benessere positivo<sup>32</sup>: una vita con benessere positivo avrà più valore quando la popolazione cui si va ad aggiungere è piccola, meno valore quando è grande. Come conseguenza di ciò, quando l’aggiunzione riguarda una popolazione piccola, il Principio del Valore Variabile tende ad assomigliare al Principio Totale – aggiunzioni di persone con benessere positivo contribuiranno al valore della popolazione in misura più o meno pari al loro valore personale<sup>33</sup> –; mentre, quando la popolazione è già grande, assomiglia sempre più al Principio della Media – per valutare il valore dell’aggiunzione bisognerà sempre più fare riferimento alle altre vite che compongono la popolazione.

Qui analizzeremo una sola versione del Principio del Valore Variabile – che chiameremo, generalizzando per comodo, PVV. Secondo PVV aggiungendo vite con benessere positivo possiamo aumentare il valore della popolazione; tuttavia, per ogni livello di benessere medio, vi è un limite al valore che una popolazione può avere a *quel* livello di benessere, non importa quanto grande possa essere la popolazione. Una volta approssimatasi al proprio valore limite – collegato al benessere medio –, che la popolazione cresca ulteriormente dieci, cento o mille volte non farà alcuna differenza: nessuna aggiunzione di vite a quel livello di benessere gli permetterà di superare il limite.

Maggiore è il benessere medio di una popolazione, maggiore sarà il valore limite che potrà raggiungere. Se immaginiamo due livelli di benessere molto diversi, uno molto alto e l’altro molto basso, e due popolazioni tanto grandi da approssimarsi al limite del valore che ciascuna può avere in considerazione del proprio livello di benessere medio, la popolazione con alto livello di benessere prossima al suo limite avrà un valore superiore rispetto a quella con basso livello di benessere, anche se quest’ultima sia numericamente molto maggiore. Dunque, data una popolazione con alto livello di benessere medio e una con basso livello di benessere medio, la seconda potrà essere migliore della prima solo qualora la prima sia lontana dal proprio limite: in quest’ultimo caso, infatti, la presenza, nella seconda popolazione, di molte più vite con

<sup>29</sup> ARRHENIUS 2000b, 7; BROOME 2004, 208.

<sup>30</sup> CARLSON 1998, 290.

<sup>31</sup> GREAVES 2017, 4.

<sup>32</sup> La precisazione è tratta da ARRHENIUS et al. 2017. È possibile concepire versioni del Principio del Valore Variabile che attribuiscono valore marginale decrescente anche al benessere negativo. Non prenderò in esame il caso. Sull’inopportunità di attribuire valore marginale decrescente anche alle aggiunzioni di vite con benessere negativo, CARLSON 1998, 295.

<sup>33</sup> Hurka delinea un Principio del Valore Variabile in cui l’aggiunzione di una vita a una popolazione molto piccola ha valore maggiore di quella che gli assegnerebbe persino il Principio Totale (HURKA 1983, 502).

benessere positivo può essere decisiva. Allora, per PVV, A è meglio di Z. Infatti, essendo A una popolazione di miliardi di persone, il suo limite può dirsi approssimato<sup>34</sup>.

PVV viola dunque il Principio della Mera Addizione. Se abbiamo una popolazione con livello di benessere alto già prossima al suo limite (A) e la compariamo a una popolazione alternativa data dalla mera addizione di un certo numero di persone con benessere più basso (A+), avendo A+ un benessere medio inferiore, A+ sarà peggiore di A. In un caso del genere, una mera addizione fa precipitare il benessere medio e il limite del valore della popolazione. L'intento di PVV è però proprio quello di *circoscrivere* la violazione del Principio di Mera Addizione ai casi più accettabili: la violazione del Principio di Mera Addizione avviene solo quando si aggiungono ulteriori persone a popolazioni già molto grandi; secondo PVV, se immaginiamo una popolazione molto piccola, un aumento della popolazione sarebbe auspicabile anche se accompagnato da una contrazione del livello di benessere medio<sup>35</sup>.

Gli esiti contro-intuitivi di PVV non si arrestano alla violazione (contenuta) del Principio della Mera Addizione. Problematica potrebbe risultare l'individuazione dei vari valori limite – e dei coefficienti di decrescita del valore marginale – di ciascuna popolazione. Sarà mai possibile fare ricorso a criteri convincenti?<sup>36</sup> Tuttavia, ancora prima che questo problema possa prendere corpo, se ne deve notare un altro. PVV implica la Conclusione Sadica.

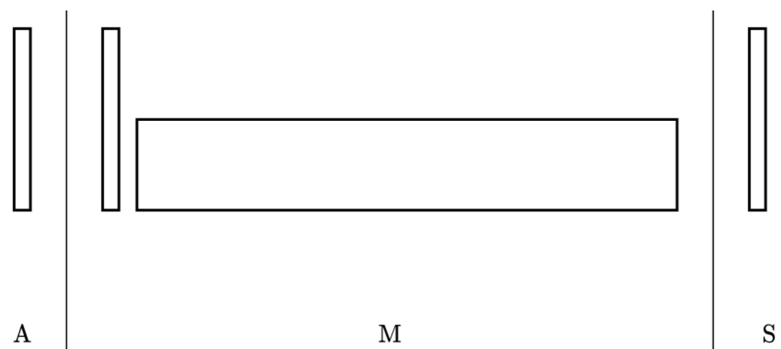


Fig. 4

<sup>34</sup> GREAVES 2017, 4. In altri termini, mentre il valore di una popolazione nel Principio della Media è pari al livello di benessere medio (B) e per il Principio Totale è pari al prodotto del benessere medio per il numero delle persone esistenti ( $nB$ ), nel PVV il valore di una popolazione è dato dalla funzione  $f(n)B$ , per cui il valore di uno scenario non viene aumentato quando – in prossimità del limite che una popolazione con un certo livello di benessere avrebbe, ossia nel caso di popolazioni già molto grandi – si aggiungessero vite che aumentano il benessere totale ma diminuiscono il benessere medio, ARRHENIUS 2000b, 60. Questa versione del Principio del Valore Variabile è stata elaborata da NG 1989, sebbene l'autore tenga a precisare che a suo avviso non costituisce una soluzione al Paradosso della Mera Addizione – viola, infatti, come vedremo subito il Principio di Mera Addizione –, quanto ciò che di più vicino possa esserci a una soluzione qualora non si voglia semplicemente accettare la Conclusione Ripugnante, NG 1989, 242-244, 249. Un principio molto simile a quello elaborato da Ng è stato difeso in HURKA 1983. TEMKIN (2012, 328 ss.) delinea una concezione simile al PVV in rapporto non alla mera aggregazione del benessere, quanto alla aggregazione e comparazione di molteplici ideali nel contesto di un'etica pluralista.

<sup>35</sup> NG 1989, 249. HURKA (1983, 502) ritiene che il PVV sia il principio che meglio dà conto del dovere che gli eventuali superstiti di una catastrofe planetaria avrebbero di aumentare la popolazione umana.

<sup>36</sup> Qui bisogna distinguere tra il limite del valore di una popolazione – relativamente al livello di benessere medio – ed il limite delle persone che compongono quella popolazione. Mentre PVV prevede un limite al valore della popolazione, non prevede un limite al numero delle persone con benessere positivo che compongono una data popolazione. Il limite può funzionare in modi diversi. Nella versione da me riportata il limite non viene mai raggiunto ma solo approssimato: l'aggiunzione di vite, che non modificano il livello di benessere medio della popolazione (altrimenti il limite cambierebbe), perde valore ma mantiene sempre un valore positivo; tuttavia il valore va diminuendo così rapidamente in prossimità del limite che il limite non può mai essere raggiunto. Secondo questa versione, aggiunzioni di vite con benessere positivo in prossimità del limite sono pressoché irrilevanti. Possiamo immaginare un'altra versione – più semplice –, in cui il limite viene raggiunto e l'aggiunzione di ulteriori persone (che, sempre, non modificano il benessere medio della popolazione) ha perso del tutto il suo valore. In entrambi i casi, comunque, il limite – ed il connesso problema della sua arbitrarietà – attiene al valore della popolazione e non alla popolazione stessa.

Fig. 4

In A – dove ciascuno ha un livello di benessere molto alto – vivono già miliardi di persone. Dato che il limite rispetto a quel livello di benessere medio è stato raggiunto, aggiunzioni di ulteriori persone con benessere – positivo, ma – inferiore alla media si tradurranno in un decremento del valore della popolazione di partenza. Può allora darsi un numero di persone aggiuntive con benessere positivo che andrebbero a diminuire il valore di A di più di quanto non lo farebbe l’aggiunzione di un minor numero di persone con benessere negativo. M è peggio di S. In altri termini, PVV ammette che sarebbe meglio, in certi casi, aggiungere vite tormentate – vite degne di non essere vissute – che aggiungere un numero maggiore di vite con benessere positivo. Una conclusione contro-intuitiva<sup>37</sup>.

#### 4. Il Principio del Livello Critico

Secondo il Principio del Livello Critico (PLC) una vita ha valore contributivo positivo se il benessere contenuto in quella vita supera un certo livello – il livello critico appunto – che è maggiore del livello zero – il livello in cui le componenti positive e negative del benessere si compensano<sup>38</sup>. PLC riflette il principio dell’Invarianza del Valore Contributivo – diversamente dal Principio della Media e dal PVV, similmente al Principio Totale – e non implica la Conclusione Ripugnante<sup>39</sup>. Z è un mondo in cui ciascuno conduce una vita che è appena degna di essere vissuta; una vita del genere è certamente sotto il livello critico e dunque non contribuisce positivamente al valore della popolazione – Z non ha valore positivo. Invece, A è un mondo in cui la qualità della vita è molto alta; chi vive in A ha una vita sopra il livello critico ed esistendo contribuisce positivamente al valore della popolazione – A ha valore positivo. Diverse versioni del Principio del Livello Critico assegnano un diverso valore alle vite sotto il livello critico.

Qualora alle vite sotto il livello critico assegniamo valore pari a zero – valore neutrale – incorriamo in una violazione del Principio di Non-Antieguagliarismo: se consideriamo una popolazione i cui membri hanno vite appena sotto il livello critico, sarebbe meglio che esistesse una popolazione di egual numero in cui un singolo individuo abbia un livello di benessere appena sopra il livello critico e tutti gli altri abbiano vite con un livello di benessere pari a zero<sup>40</sup>.

Se alle vite sotto il livello critico attribuiamo, invece, valore negativo<sup>41</sup> – per cui il valore contributivo di ciascuna vita sarà pari alla differenza tra il livello critico e il valore personale di quella vita –, PLC non violerà il Principio di Non-Antieguagliarismo. Questa versione di PLC costituisce una deroga al Principio di Mera Addizione: alcune mere addizioni aumenteranno il valore della popolazione, quelle sopra il livello critico, altre – quelle sotto il livello critico – lo diminuiranno. Sebbene questa soluzione possa sembrare a prima vista un compromesso accettabile, essa ingenera la Conclusione Sadica: se, infatti, l’aggiunzione di vite con benessere positivo ma inferiore al livello critico ha valore contributivo negativo, allora per ogni mondo in cui esistono solo persone con vite tormentate – vite con benessere largamente negativo – si dà un mondo in cui esistono molte più persone con benessere positivo ma inferiore al livello critico, tale che il secondo mondo è peggiore del primo<sup>42</sup>. Per contenere

<sup>37</sup> ARRHENIUS 2000b, 63 ss.; CARLSON (1998, 301 ss.) difende la Conclusione Sadica in certi casi. SIDER (1991) ha difeso una versione del Principio del Valore Variabile – quella qua analizzata è infatti solo una versione del Principio del Valore Variabile – che non va incontro alla Conclusione Sadica, ma questa versione dà esiti contrari al Principio di Non-Antieguagliarismo (ARRHENIUS 2000b, 67-71).

<sup>38</sup> ARRHENIUS et al. 2017.

<sup>39</sup> CARLSON 1998, 292; GREAVES 2017, 6

<sup>40</sup> Così CARLSON 1998, 292; questa versione del Principio del Livello Critico è discussa in PARFIT (1984, 412 s.) e criticata su altre basi.

<sup>41</sup> Così BLACKORBY, BOSSERT, DONALDSON 1997, 202; vedi anche BROOME 2004, 209-212.

<sup>42</sup> ARRHENIUS 2000b, 73.

il problema posto dalla Conclusione Sadica può ipotizzarsi che il livello critico sia abbastanza basso – prossimo al livello in cui il benessere ha valore zero. Tuttavia, così congegnandolo, si espone PLC sostanzialmente alla Conclusione Ripugnante. In generale: più alziamo il livello critico, più vite con benessere positivo avranno valore contributivo negativo, più ci terremo lontani dalla Conclusione Ripugnante, ma anche via via peggiori saranno le istanze della Conclusione Sadica che verranno implicate; e, per converso, più teniamo basso il livello critico – vicino al livello zero – meno la Conclusione Sadica costituirà un problema, ma più ci avvicineremo alla Conclusione Ripugnante<sup>43</sup>.

Secondo una terza versione di PLC il livello critico non sarebbe unico, ma costituito da una banda di livelli di benessere con il livello zero come limite inferiore e un livello  $x > 0$  come limite superiore. Nel momento in cui vengono confrontate due popolazioni di diversa grandezza, una popolazione potrà dirsi migliore dell'altra solo se risulta migliore per ognuno dei livelli critici che compongono la banda<sup>44</sup>. Qualora però una popolazione sia migliore di un'altra rispetto ad alcuni dei livelli critici della banda, ma peggiore rispetto agli altri, allora la relazione assiologica tra le due popolazioni sarà *vaga, incommensurabile, imprecisa* – non distinguo tra questi casi<sup>45</sup>.

Nemmeno questa terza versione di PLC rappresenta una soluzione soddisfacente. Compariamo A e Z (Fig. 1). Dato che, per certi livelli critici della banda, A è migliore di Z, ma, per altri, Z è migliore di A – ricordiamo infatti che la banda di livelli critici ha come limite inferiore zero, e che dunque vi sono livelli critici tanto bassi da attribuire un maggiore valore a Z –, seguirà che A e Z sono incommensurabili, o intrattengono una relazione imprecisa, vaga. Avremo una forma attenuata della Conclusione Ripugnante. Lo stesso dicasi per la comparazione tra M e S (Fig. 4). Anche in questo caso vi saranno livelli critici in cui M è migliore di S, ma, se il limite superiore della banda è abbastanza alto, vi saranno anche livelli critici in cui S è migliore di M. Le due popolazioni saranno allora incommensurabili o intratterranno una relazione imprecisa, vaga – avremo una versione attenuata della Conclusione Sadica. La banda di livelli critici crea una voragine che ingoia molti dei nostri giudizi assiologici<sup>46</sup>.

## 5. Il Principio Lessicale

Archiviato PLC, consideriamo una terza ipotesi risolutiva, difesa da Parfit fino ai suoi ultimi anni. Secondo il Principio Lessicale (PLE), tra le vite con benessere positivo, dobbiamo distinguere vite

<sup>43</sup> BROOME 2004, 213.

<sup>44</sup> A tal proposito varranno le regole appena viste per la seconda versione del PLC: per stabilire se una popolazione sia migliore dell'altra dato un certo livello critico, si calcherà il valore contributivo di ciascun membro delle popolazioni determinando la differenza (positiva o negativa) tra il suo valore personale ed il livello critico. Ragioneremo come se quel livello critico fosse l'unico, per poi passare al successivo e fare altrettanto.

<sup>45</sup> BLACKORBY et al. 1997, 216 ss. (vedi anche ARRHENIUS 2000, 74 ss.). Per fare un esempio, ipotizziamo che la banda dei livelli vada da 0 a 10, abbiamo due popolazioni: A (4,4,...,4,4,Ω) e B (4,4,...,4,2,4) dove ogni numero corrisponde a milioni di individui a un determinato livello di benessere – 4 o 2 –, e la Ω indica l'assenza di individui. A è dunque una popolazione inferiore per numero rispetto a B, ma con una qualità media della vita più alta. Per chiederci se A sia migliore di B, dobbiamo chiederci se A sia migliore di B qualora il livello critico fosse 1, poi qualora fosse 2, poi 3 e così via fino al livello 10. Emerge che A è migliore di B solo se il livello critico è superiore a 2: infatti per i valori (v) superiori a 2 sarà vero che  $(4-v) > (2-v) + (4-v) -$  si noti che per raffrontare A e B, basterà comparare le ultime due coppie di valori, essendo in ogni altro aspetto popolazioni identiche. Questo vuol dire che A è meglio di B per tutti i livelli critici che vanno da 3 a 10, mentre B è meglio di A per i livelli che vanno da 0 a 2. Siccome nessuna delle due popolazioni è migliore dell'altra per tutti i livelli critici rilevanti, si deve concludere che le due popolazioni sono incommensurabili, o che intrattengono rapporti assiologici vaghi (BROOME 2004, 202-205).

<sup>46</sup> Così ARRHENIUS 2000, 75 s. Sulla distinzione tra incommensurabilità e vaghezza, BROOME 2004, 165 ss.; sia l'indeterminatezza (p. 170) sia la vaghezza (p. 182, 202 ss., 213) generano la voragine che ingoia fin troppi giudizi assiologici. Broome mostra anche come questa versione di PLC implichi che una popolazione maggiore possa essere peggiore di una popolazione minore con un livello di benessere pure minore: un cataclisma che riducesse la popolazione mondiale e diminuisse il benessere di alcuni dei superstiti sarebbe un evento auspicabile (si rimanda per la dimostrazione a BROOME 2004, 204).

eccelse e vite non-eccelse: per quanto l'aggiunzione di vite non-eccelse abbia sempre valore positivo e per quanto alto possa essere il valore aggregato delle vite non-eccelse, il valore delle vite non-eccelse non potrà mai superare il valore di un qualsiasi numero – o di un numero qualificato – di vite eccelse<sup>47</sup>. Notiamo come, seguendo PLE, *qualsiasi* vita degna di essere vissuta, anche una vita non-eccelsa, contribuisce positivamente a conferire valore a una data popolazione – in ciò la differenza rispetto a PLC – e il suo contributo non diminuisce, né altrimenti varia, in ragione del benessere degli altri componenti della popolazione. Non vi è un limite al valore di una popolazione composta da vite non-eccelse – in ciò la differenza rispetto a PVV e al Principio della Media.

Secondo PLE, A sarebbe meglio di Z: infatti, A è composto solo da vite eccelse, mentre nessuna vita di Z è una vita eccelsa. In questo senso PLE evita la Conclusione Ripugnante. Tuttavia, PLE viola platealmente il Principio di Non-Antiegalitarismo: dato il mondo A+ in cui una certa minoranza di persone ha vite eccelse, mentre la stragrande maggioranza ha vite appena degne di essere vissute (Fig.3.), questo mondo sarebbe migliore – per via del valore della minoranza eccelsa – di un mondo Z, composto da persone con un livello di benessere molto basso – sebbene superiore a quello dei più svantaggiati di A+ – in cui il benessere medio e totale è maggiore di quello di A+<sup>48</sup>.

Se si ragiona solo in termini di benessere e a fare la differenza tra una vita eccelsa e una non-eccelsa è soltanto la quantità di benessere, PLE appare anti-egualitario, elitario e davvero contro-intuitivo. Una difesa di PLC presuppone il riconoscimento di valori diversi dal benessere: valori di carattere estetico, come il raggiungimento dell'eccellenza artistica, ma anche la sperimentazione di relazioni affettive sofisticate – *Perfezionismo*. Si dirà che A e Z non differiscono solo per il numero di persone esistenti e il loro livello di benessere. Si specificherà che A è un mondo in cui esiste più arte, più acume, più purezza nelle amicizie, più romanticismo nelle relazioni sentimentali – più cose belle –, mentre Z è un mondo in cui la vita, riassumendola con un'immagine, è musica da sala d'attesa e patate (“muzak and potatoes”<sup>49</sup>).

Si noti, questa mossa equivale a cambiare i termini del problema del Paradosso della Mera Addizione come finora posti. In si abbiamo detto che le popolazioni comparate – in primo luogo, A e Z – presentano una diversa dimensione e un diverso livello di benessere, mentre altre caratteristiche di queste due popolazioni sono tenute costanti e non hanno, dunque, rilevanza – clausola *ceteris paribus*. Invocare la perdita di valori perfezionisti nel passaggio da A a Z equivale a scorporare questi valori dalla clausola *ceteris paribus* – A e Z sono diverse *anche* perché in Z ci sono molte meno cose belle di quante ve ne siano in A. Questa rappresentazione di A e Z non è necessaria: potremmo immaginare che in Z esistano le stesse cose belle che esistono in A – ma-

<sup>47</sup> PARFIT 1984, 413 ss. Secondo questa qualificazione del PLE non esiste un numero di vite non-eccelse che superi il valore contributivo di una singola vita eccelsa. PLE è il principio che, in suoi successivi lavori, Parfit crede possa offrire una soluzione ai dilemmi dell'etica della popolazione. In questi suoi successivi lavori (PARFIT 1986, 162; PARFIT 2016, 112), Parfit adotta in relazione al PLE la clausola per cui il mondo A sarebbe comunque composto da miliardi di persone. Secondo questa ultima versione del PLE, non esiste un numero di vite non-eccelse che superi il valore di miliardi di vite eccelse.

<sup>48</sup> In realtà, PLE ha conseguenze ancora più inegualitarie della mera violazione del Principio di Non-Antiegalitarismo. Una posizione è semplicemente anti-egualitaria qualora ammette che, comparando due popolazioni di eguali dimensioni, si diano *certi* casi in cui la popolazione (i) in cui i più avvantaggiati stanno meglio che nell'altra – in cui i “picchi” di benessere sono più alti – e (ii) con benessere medio (e totale) inferiore sia la popolazione migliore. Un anti-egualitario può però ammettere che in *certi* altri casi la popolazione con benessere medio e totale inferiore ma con più alti picchi sia peggiore dell'altra. Un anti-egualitario può dunque accettare che vi sia un incremento nel benessere di un certo numero di persone che compensi un minimo decremento nel benessere di una persona. PLE va molto oltre: se immaginiamo una prima popolazione, composta da una singola vita eccelsa e da miliardi di altre vite con benessere molto basso, ed una seconda popolazione di egual misura, composta da vite che hanno *quasi* lo stesso benessere della persona con vita eccelsa ma che non sono – per pochissimo – vite eccelse, la prima popolazione sarebbe preferibile alla seconda. In altri termini, PLE nega che vi sia un incremento nel benessere di vite non-eccelse – e che rimangono, per poco, non-eccelse anche dopo l'incremento – che possa compensare il decremento, anche minimo, di, al limite, una sola vita eccelsa, qualora questo decremento renda quella vita non più eccelsa (ARRHENIUS 2000b, 97-100).

<sup>49</sup> PARFIT 1986, 162

gari che ne esistano anche di più – ma che ciascuno non ne ricavi grande benessere – ad esempio, perché ciascuno in Z vive una vita che è molto breve. La difesa di PLE a mezzo di un’ottica perfezionista funziona *a patto* che intendiamo A come un mondo dall’elevato valore estetico e Z un mondo in cui questo valore è praticamente nullo<sup>50</sup>.

Un primo problema riguarda PLE e il valore del Perfezionismo – un difetto che potrebbe essere colto anche qualora i valori perfezionisti fossero gli unici esistenti, Monismo Perfezionista. A una categoria di fatti può riconoscersi valore lessicalmente superiore a un’altra quando si tratta di categorie molto diverse – il peccato e il dolore sono entrambi dei mali, ma il peccato è più grave di qualsiasi quantità di dolore<sup>51</sup>. Tuttavia, una differenza lessicale non sembra appropriata all’interno della medesima categoria. Immaginiamo una vita piena di cose belle. È difficile credere che questa vita abbia un valore lessicalmente superiore a qualsiasi numero di vite in cui quasi tutte le stesse cose belle sono presenti<sup>52</sup>.

In risposta a questo problema si invoca il concetto di *imprecisione*. Davanti a certe alternative, come due diverse carriere, sembra assurdo dire che l’una sia migliore, ad esempio, 3,7 volte dell’altra: vi sono casi in cui le uniche verità nel rapporto tra due alternative sono imprecise. Quando due cose sono imprecisamente l’una “non peggiore” dell’altra – termine tecnico volto a denotare l’imprecisione –, esiste un terzo termine di comparazione che, sebbene abbia una certa relazione precisa con uno dei primi termini – sia o migliore o peggiore o uguale –, non intrattiene la medesima relazione con l’altro – una deroga alla transitività. L’essere uno scrittore di successo non è peggio del fare carriera come medico, e nemmeno l’essere uno scrittore di successo appena minore è peggio del fare carriera come medico, però l’essere uno scrittore di successo minore è peggio di essere uno scrittore di successo appena maggiore.

Potremmo allora ragionare come di seguito. Data una prima popolazione composta da un certo numero di vite ciascuna piena di beni perfezionistici, possiamo immaginare una seconda popolazione più grande della prima in cui la vita di ciascuno contiene quasi gli stessi beni perfezionistici, tale che la prima popolazione non sarà lessicalmente migliore della seconda, ma solo non peggiore di questa. Se però immaginiamo una terza popolazione maggiore della seconda, in cui la vita di ciascuno contiene ancora meno beni perfezionistici, potrà essere vero che questa terza popolazione, che è solo non peggiore della seconda, sia lessicalmente peggiore rispetto alla prima – eviteremo così la Conclusione Ripugnante<sup>53</sup>.

Un secondo problema per PLE deriva dal pluralismo etico – questo problema, a differenza del precedente, non riguarderebbe coloro che adottano una forma di Monismo Perfezionista. Prendiamo in esame la seguente alternativa

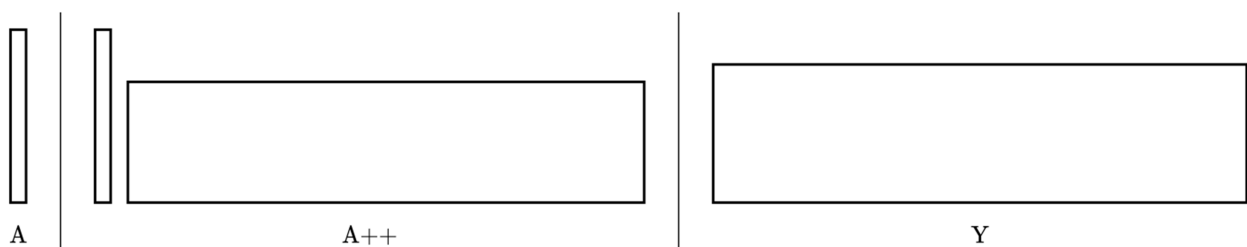


Fig. 5

<sup>50</sup> PARFIT 2016, 117; PARFIT 1986, 162.

<sup>51</sup> Questa gerarchia dal sapore quantomai inattuale si attribuisce al Cardinale Newman e si adopera come esempio classico della lessicalità (PARFIT 1984, 413 s.).

<sup>52</sup> PARFIT 1984, 414; PARFIT 2016, 116.

<sup>53</sup> PARFIT 2016, 113-121.



Ammettiamo che A++ sia almeno buono tanto quanto A – Principio della Mera Addizione –<sup>54</sup>, se riteniamo anche – Principio di Non-Antieguaglianza – che Y sia migliore di A++, si dischiude la strada verso la Conclusione Ripugnante – possiamo reiterare le mere addizioni e le redistribuzioni del benessere finché, per transitività, non otterremo che Z sia migliore di A. Chiediamoci se Y non possa essere migliore di A++. Nel contesto di un’etica pluralista, Y è migliore di A++ per quanto concerne il benessere totale, l’eguaglianza e il Maximin, mentre dal punto di vista del Perfezionismo le cose non sono chiare. Nel passaggio da A++ a Y molte delle cose belle della vita andrebbero perse, posto che nessuno raggiungerebbe le eccellenze che contraddistinguono i più fortunati di A++, però molte delle cose belle rimarrebbero. Se dire che Y sia meglio di A++ (tutto considerato) sembra troppo, una soluzione accettabile potrebbe essere quella di dire che Y è non peggiore di A++. Questo basterebbe a evitare la Conclusione Ripugnante. Dal momento che le relazioni assiologiche imprecise derogano alla transitività, pur concedendo che Y sia non peggiore di A++ potremmo sostenere che Z – un mondo in cui tutte (non solo alcune) le cose belle sono andate perdute – sia lessicalmente peggiore di A++.

Ma, in radice, è plausibile sostenere che Y sia solo imprecisamente non peggiore di A++ e non, invece, che Y sia meglio di A++? Parfit a riguardo ammette che *in sé* non è plausibile sostenere questo rapporto, però è meno implausibile dell’acceptare la Conclusione Ripugnante<sup>55</sup>. Tralasciando la rilevanza del criterio indicato – la soluzione “meno implausibile”<sup>56</sup> –, notiamo come l’instabilità del PLE sia resa ancora più evidente se consideriamo, in luogo del precedente, il seguente caso

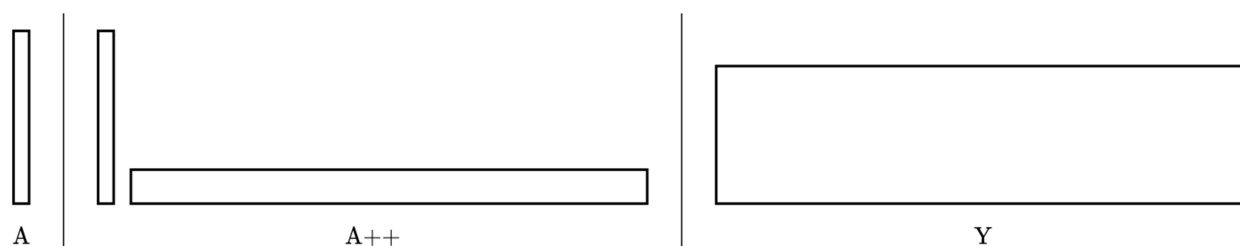


Fig.6

In questa versione il vantaggio di Y su A++ rispetto ai valori del benessere, dell’eguaglianza e del Maximin è molto maggiore del caso analizzato da Parfit. Sembra ancora più plausibile che Y sia semplicemente migliore di A++ e non soltanto imprecisamente non peggiore di A++. Eppure per sostenere che A++ sia migliore di Z ed evitare così la Conclusione Ripugnante, dovremmo sostenere che Y sia semplicemente non peggiore di A++ anche in questa versione.

## 6. “Person-affecting moralities”. *La restrizione personalista*

Le precedenti ipotesi risolutive avevano come presupposto che fosse possibile stabilire relazioni assiologiche tra popolazioni alternative solo guardando alla quantità di benessere, alla sua distribuzione ed, eventualmente, all’esistenza delle cose belle della vita. Sotto l’etichetta di “person-

<sup>54</sup> Parfit rafforza il problema sotto questo aspetto, ponendo l’alternativa tra A e un mondo simile ad A++ in cui però il benessere di coloro che sono più fortunati è addirittura più alto di quello di coloro che vivono in A – “Raised A+”. Questo porta Parfit ad assumere che “Raised A+” sia migliore di A (PARFIT 2016,124). Parfit probabilmente ha voluto rendere così più chiaro come non basti – per evitare la Conclusione Ripugnante – negare il Principio della Mera Addizione. Per non complicare in questo punto l’esposizione – posto che non si tratta di un aspetto cruciale, operando PLE nelle relazioni tra A++ e B e non in quella tra A++ e A – ho preferito modificare l’esempio di Parfit.

<sup>55</sup> PARFIT 2016, 124-127.

<sup>56</sup> Tornerò sul punto alla nt. 103.

*affecting moralities*”, invece, si è soliti ricomprendere una serie di principi o teorie accomunati dall’assegnare importanza a fattori diversi da quelli sopra elencati – ad esempio, l’essere quella persona esistente nel presente o esistente nel futuro, oppure l’esistere in tutti o solo in alcuni degli scenari comparati. Sebbene assegnino importanza a fattori ulteriori rispetto a quelli considerati nei precedenti paragrafi, le teorie *person-affecting* devono pure utilizzare dei principi per aggregare il benessere e questi possono essere simili a quelli visti precedentemente. Dunque, una volta attribuito valore a un certo fattore – per esempio, l’essere una certa popolazione composta da persone presenti anziché future –, potremo distinguere una versione che aggrega il valore attraverso il Principio Totale, un’altra attraverso il Principio del Livello Critico, e così via. Ciò rende possibile che i problemi di questi principi vengano ereditati dalla teoria *person-affecting* che li incorpora<sup>57</sup>. La speranza dell’approccio *person-affecting* è precisamente che il riconoscimento di rilevanza ai fattori aggiuntivi risolva i problemi dei principi aggregativi. In questo paragrafo prenderò in esame la versione *person-affecting* del Principio Totale.

La famiglia di teorie *person-affecting* sulla quale ci concentreremo distingue tra persone necessarie – ossia persone che esistono in tutte le popolazioni comparate – e persone meramente possibili – persone che esistono in un solo scenario, oppure (ma ci disinteresseremo di questo caso) in alcuni ma non tutti –, e attribuisce rilevanza esclusiva – versione estrema –, rilevanza particolare – Estinzionismo, versione asimmetrica, versione condizionale –, o semplicemente maggiore – versione moderata – al benessere delle prime. Chiamiamo questa famiglia di teorie “Comparativismo”<sup>58</sup>. Immaginiamo allora il caso “curare o rimpiazzare” (§1):

*Curare o rimpiazzare*

*Scelta 1:* Tizio ha benessere 80, Caio non esiste.

*Scelta 2:* Tizio ha benessere 30, Caio esiste e ha benessere 90.

Possiamo incrementare il benessere di una persona esistente – prolungare la sua vita degna di essere vissuta, alleviarne le sofferenze, darle più momenti di gioia –, o mettere al mondo una persona aggiuntiva che vivrà una vita con benessere positivo. Se effettuiamo la seconda scelta, il benessere totale sarà notevolmente maggiore di quello che avremmo ove effettuassimo la prima scelta. L’Utilitarismo Classico – PUT – stabilisce che la seconda scelta è quella migliore (§1).

Versioni diverse del Comparativismo danno risposte diverse a questa scelta.

Secondo la versione moderata del Comparativismo – *Comparativismo Morbido*<sup>59</sup> – il benessere di tutte le persone presenti nelle popolazioni alternative deve essere preso in considerazione, ma

<sup>57</sup> ARRHENIUS 2009, 306.

<sup>58</sup> Il termine “Comparativismo” è tratto da ARRHENIUS 2009, 293. Tuttavia, nella terminologia di Arrhenius il Comparativismo distingue tra persone che esistono in entrambi gli scenari e persone che esistono in uno solo degli scenari considerati – ciò che chiamerò “Comparativismo” corrisponde a quello che Arrhenius chiama “Necessarismo”. A ogni modo, negli esempi che seguiranno non porrò il caso che la stessa persona possibile esista in più di uno scenario. Quando parleremo di persone possibili, in altri termini, parleremo di persone che esistono in uno solo degli scenari comparati – dunque il mio uso del termine “Comparativismo” non si differenzierà, nei fatti, dall’uso che ne fa Arrhenius. Si tratta di una scelta che vuole rispondere all’esigenza di non complicare troppo un’introduzione al dibattito. Un’altra teoria *person-affecting* distingue tra persone vive nel presente e persone future ed assegna carattere decisivo – versione estrema – o, semplicemente, riconosce maggior valore – versione moderata – al benessere delle prime. Ad esempio (PARFIT 1984, 361-364), possiamo immaginare una prima popolazione, composta da una generazione presente (noi) e molte generazioni future, in cui la generazione presente ha un alto livello di benessere, ed una seconda popolazione alternativa in cui la generazione presente ha benessere più basso ma le generazioni future hanno un livello di benessere ben più alto di quello che avrebbero nella prima popolazione. Assegnare carattere dirimente al maggior benessere delle persone presenti vorrebbe dire che la prima popolazione sia migliore della seconda, ad onta del fatto che contiene minor benessere totale e c’è più ineguaglianza intergenerazionale. La distinzione tra persone future e persone possibili riveste una certa importanza in etica della popolazione, tuttavia per esigenze espositive non mi ci soffermerò.

<sup>59</sup> ARRHENIUS 2009, 301 ss.

– diversamente da PUT – dobbiamo dare maggiore peso al benessere delle persone che esistono in entrambi gli scenari comparati. Il valore contributivo del benessere delle persone che, come Caio, esistono in una sola delle alternative comparate è sottoposto a un tasso di sconto, più o meno elevato – un problema interno al Comparativismo Morbido. A seconda del tasso prescelto, la seconda scelta sarà preferibile o meno rispetto alla prima. In ogni caso, anche se optiamo per una versione del Comparativismo Morbido che, applicando un tasso di sconto sufficientemente elevato, giustificherebbe la prima scelta, si darà sempre una scelta tra il curare una persona malata e il mettere al mondo una certa quantità di persone aggiuntive con benessere positivo che sarà preferibile alla prima scelta. Da ciò segue che il Comparativismo Morbido – nella misura in cui adotta come principio aggregativo il Principio Totale – ingenera la Conclusione Ripugnante.

Secondo le forme di Comparativismo che analizzeremo di seguito, dal momento che l'unica persona necessaria è Tizio, e Tizio sta meglio se adottiamo la prima politica, allora la prima politica è quella migliore. Per l'Utilitarismo Classico potremmo avere ragioni per lasciare morire molte persone che giacciono negli ospedali al fine di dirottare le nostre risorse verso la creazione di un maggior numero di persone sane e forti – pensiamo a energie politiche a favore della natalità. Così anche per il Comparativismo Morbido. Per le forme di Comparativismo che verranno analizzate di seguito, invece, dobbiamo occuparci di rendere le persone felici e non di creare nuove persone felici – questa la fortunata idea di base, le cui radici risalgono agli albori dell'etica della popolazione contemporanea<sup>60</sup>.

Iniziamo. Secondo il *Comparativismo Estremo* le uniche vite che hanno valore contributivo sono le vite delle persone che esistono in entrambi gli scenari comparati: questo valore sarà positivo, se il benessere avrà somma netta positiva, e negativo, se avrà somma netta negativa. Dovremmo, in altre parole, disinteressarci del tutto della condizione delle persone che esistono in solo uno degli scenari comparati<sup>61</sup>.

Il Comparativismo Estremo, nell'attribuire valore contributivo alle sole vite delle persone che esistono in entrambi gli scenari comparati, assume la concezione *tradizionale* o *comparativa* del danno e del beneficio. Secondo la concezione comparativa, danni e benefici sono eventi che fanno stare *qualcuno* rispettivamente peggio o meglio di come sarebbe stato qualora l'evento non si fosse verificato, e questo presuppone che quella persona esista o esisterà, indipendentemente dal se l'evento dannoso o benefico si produca. In altri termini, per la concezione comparativa del danno e del beneficio, gli unici eventi dannosi o benefici per una persona sono quegli eventi che non determinano essi stessi l'esistenza di quella persona, ma che, presupposta questa esistenza, incidono sul suo benessere<sup>62</sup>.

Secondo il Comparativismo Estremo gli unici danni e benefici di cui si deve tenere conto – gli unici che hanno un valore contributivo, rispettivamente, negativo e positivo – sono danni e

<sup>60</sup> Lo slogan recita «We are in favor of making people happy, but neutral about making happy people» (NARVESON, 1973, 80). Sull'essere neutrali rispetto alla creazione di nuove persone si veda anche BROOME 2004, 143 ss. Notiamo come anche muovendo da questa idea possono essere giustificate politiche che sfavoriscono persone esistenti e favoriscono la nascita di persone aggiuntive che vivranno vite degne di essere vissute. Ad esempio, se ipotizziamo che le politiche in sostegno delle nascite incrementeranno il benessere dei genitori in misura maggiore di quanto decremeranno il benessere dei non-genitori, valuteremo come giuste queste politiche anche se quel che ci sta a cuore è “il rendere le persone felici e non il creare persone felici”. Ma si noti che il motivo per cui sarebbero giuste – per il Comparativismo – risiede sempre nel trattamento di persone esistenti – i possibili genitori – e non nel benessere delle persone possibili – gli eventuali figli. Il valore delle nascite aggiuntive è *strumentale* al benessere delle persone esistenti.

<sup>61</sup> ARRHENIUS 2009, 293. Qui di seguito farò largo uso degli scritti di Jeff McMahan sulla nozione di benefico e danno nell'ambito dell'etica della popolazione. La tassonomia di Arrhenius verrà dunque analizzata sulla base delle nozioni introdotte da McMahan.

<sup>62</sup> «[...] ordinary benefits and harms, which are bestowed or inflicted on existing or future people, whose existence is independent of the act that causes or constitutes the benefit or harm», MCMAHAN 2013, 7.

benefici comparativi<sup>63</sup>. Poiché le uniche vite in cui trovano spazio danni e benefici comparativi sono le vite di coloro che esistono in entrambi gli scenari comparati, allora solo queste avranno un valore contributivo.

Il Comparativismo Estremo ha conseguenze altamente contro-intuitive. Immaginiamo il caso “gioia o tormento”:

*Gioia o tormento*<sup>64</sup>

*Scelta 1:* Tizio e Caio avranno un livello di benessere molto alto (una vita gioiosa), Sempronio non esisterà.

*Scelta 2:* Tizio avrà un livello di benessere molto alto (una vita gioiosa), Caio non esisterà e Sempronio avrà un livello di benessere ampiamente negativo (una vita piena di tormenti).

Dal momento che l'unica persona che esiste in entrambi gli scenari, Tizio, non subisce variazioni di benessere a seconda della politica da noi scelta – non si registrano danni comparativi –, si deve concludere che le due scelte sarebbero egualmente desiderabili: mettere al mondo una persona con una vita che si prospetta gioiosa o una persona destinata a vivere tormenti infernali sarebbero scelte egualmente buone.

La concezione tradizionale, comparativa, del danno e del beneficio porta a conclusioni contro-intuitive. Ciò che il caso “gioia o tormento” mostra è che molti di noi ritengono che sia possibile danneggiare qualcuno dandogli l'esistenza, qualora la vita a cui andrà incontro sarà una vita piena di tormenti non compensati da altrettante gioie. Ammettiamo allora che, accanto ai danni comparativi, esistano dei danni *non-comparativi* o danni *esistenziali*: danni che vengono provocati dallo stesso evento che dà la vita a una certa persona. Solo *danni* esistenziali? Come quando creiamo una persona che vivrà una vita tormentata, ammettiamo che stiamo danneggiando – in senso non-comparativo, danno esistenziale – quella persona, così, quando creiamo una persona che avrà una vita in cui la somma netta del benessere è positiva, potremmo anche ammettere che stiamo elargendo un *beneficio* – anch'esso non-comparativo, beneficio esistenziale – a quella persona. In altri termini, la concezione tradizionale, comparativa, del danno e del beneficio può essere arricchita ammettendo danni e benefici esistenziali<sup>65</sup>.

Se però ammettiamo l'esistenza dei benefici esistenziali ammettiamo che creare vite con benessere positivo ha un valore che può compensare decrementi di benessere delle persone esistenti. Ossia, ammettiamo che, in casi come “curare o rimpiazzare”, sarebbe bene “rimpiazzare” e arriviamo alla Conclusione Ripugnante. Potrebbe sembrare promettente, dunque, attribuire sì rilevanza ai danni esistenziali – e così evitare che il caso “gioia o tormento” costituisca un problema – ma non anche ai benefici esistenziali – “curare” sarebbe meglio che “rimpiazzare”, la Conclusione Ripugnante sarebbe evitata. Questa concezione, tuttavia, richiede un'ulteriore precisazione.

Possiamo immaginare una vita che contiene solo piacere e una vita che contiene solo dolore. Creare una vita del primo tipo è un caso paradigmatico di beneficio esistenziale, crearne una del

<sup>63</sup> Le premesse non sono solo (i) che perché qualcuno possa essere danneggiato da un evento quel qualcuno deve stare peggio di come sarebbe stato qualora l'evento non si fosse prodotto, e (ii) che perché qualcuno possa stare peggio di come sarebbe stato qualora l'evento non si fosse prodotto quel qualcuno deve esistere anche qualora l'evento non si fosse prodotto, ma anche (iii) che perché possa esservi qualcosa di male deve esservi qualcuno che è stato danneggiato (per un esame accurato di questi e altri passaggi BOONIN 2014, 2-5).

<sup>64</sup> Seguo qua ARRHENIUS 2009, 243; vedi anche ARRHENIUS, RABINOWICZ 2015, 426 s.

<sup>65</sup> MCMAHAN 2013, 6 ss.; si veda anche MCMAHAN 2009, 50 ss. Una parte rilevante dell'indagine sul danno e beneficio esistenziale riguarda la determinazione dell'*estensione* dei medesimi. Immaginiamo una persona nata con una dolorosa malformazione dovuta alla negligenza del medico nel prescrivere una terapia per la fertilità alla madre. Rientra nell'estensione del danno esistenziale da negligenza medica solo la sofferenza derivante dalla malformazione o tutte le sofferenze indipendenti da questa – lo schiacciarsi un piede – che non si sarebbero verificate, qualora il medico non avesse prescritto l'inopportuna terapia per la fertilità, perché la persona in questione non sarebbe nata? (MCMAHAN 2013, 9). Non mi occuperò di simili questioni.

secondo tipo è un caso paradigmatico di danno esistenziale. Sebbene siano casi paradigmatici, sono anche casi irrealistici o non rappresentativi: le vite umane contengono, quasi sempre, sia piaceri sia dolori. Diciamo allora che, nella misura in cui i singoli piaceri e i singoli dolori siano riconducibili all'evento che dà la vita a una data persona, ciascuno di questi singoli piaceri corrisponde a un beneficio esistenziale e ciascuno dei singoli dolori corrisponde a un danno esistenziale. La persona avrà ricevuto *nel complesso* un beneficio esistenziale se la somma dei singoli benefici esistenziali supera la somma dei singoli danni esistenziali; e, viceversa, avrà ricevuto complessivamente un danno esistenziale se la somma dei singoli benefici esistenziali è superata dalla somma dei singoli danni esistenziali.

Ebbene, se attribuiamo rilevanza ai soli danni esistenziali ma non anche ai benefici esistenziali, otteniamo che l'aggiunzione di qualsiasi vita che conterrà almeno un dolore costituisce un danno alla persona che esisterà, l'aggiunzione sarà male – e questo indipendentemente dalla quantità di piaceri che pure saranno contenuti in quella vita. Chiunque potesse avere un figlio, avrebbe una ragione *prima facie*<sup>66</sup> per non metterlo al mondo, anche se si può ragionevolmente aspettare che suo figlio avrà una vita colma di ogni gioia. Avremmo ragioni per estinguerci – chiamiamo questa posizione *Estinzione*.

Per evitare l'Estinzione dobbiamo allora riconoscere che i benefici esistenziali abbiano *almeno* il valore di compensare i danni esistenziali – valore *compensativo* (“canceling”). Sicché, nel caso in cui le vite delle persone che potrebbero essere messe al mondo conterranno più piaceri che dolori, non sarà male creare quelle vite. Chiamiamo *Comparativismo Asimmetrico* la posizione secondo la quale non vi è un valore – e nemmeno un disvalore, in questo la differenza con l'Estinzione – nel creare persone che avranno vite con benessere positivo, ma vi è un disvalore nel creare persone con benessere negativo<sup>67</sup>. Questo principio sembra un completamento ragionevole della massima per cui dobbiamo occuparci del rendere felici le persone ma non anche del creare persone aggiuntive felici: tutto ciò è fatto salvo, ma si aggiunge che dobbiamo avere cura di non creare persone aggiuntive che avranno vite con benessere negativo.

Sebbene attraente, pure il Comparativismo Asimmetrico genera conclusioni contro-intuitive. Immaginiamo che una giovane ragazza intenda avere un figlio. Per via della sua inesperienza, della sua fragilità psicologica e della sua situazione economica suo figlio andrà incontro a notevoli difficoltà. Nel complesso, tuttavia, la vita del figlio sarà degna di essere vissuta. Se la giovane ragazza aspettasse qualche anno prima di avere un figlio, potrebbe risparmiargli queste difficoltà. Tuttavia, se aspetta di avere un figlio, a nascere sarà una persona diversa da quella che sarebbe nata qualora non avesse aspettato<sup>68</sup>. Per il Comparativismo Asimmetrico non vi è disvalore nel fatto che la ragazza non aspetti. Generalizzando, se prendiamo in esame le due scelte seguenti

*Agiatezza o stenti.*

Scelta 1: Tizio esiste e ha un elevato livello di benessere, Caio esiste e ha pure un livello di benessere elevato, Sempronio non esiste.

Scelta 2: Tizio esiste e ha un elevato livello di benessere, Caio non esiste, Sempronio esiste e ha un livello di benessere positivo ma basso<sup>69</sup>.

per il Comparativismo Asimmetrico non vi è differenza tra queste due scelte.

<sup>66</sup> La precisazione è dovuta alla circostanza che questa impostazione non esclude – similmente a quanto visto nella nt. 60 – il valore *strumentale* della nascita del figlio, il valore collegato all'incremento di benessere nella persona che desidera avere un figlio e dà attuazione a questo suo desiderio, MCMAHAN 2012, 22.

<sup>67</sup> MCMAHAN 2013, 16. McMahan elabora l'Asimmetria in MCMAHAN 1981. Per il Comparativismo Asimmetrico si veda ARRHENIUS 2009, 294.

<sup>68</sup> PARFIT 1984, 358. Si tratta di una formulazione del c.d. *Non-Identity Problem* basata sul caso delle ragazze-madri.

<sup>69</sup> MCMAHAN 2013, 6; ARRHENIUS 2009, 294 s.; PARFIT 2017, 120.

Andiamo avanti. Nell'ambito dei benefici esistenziali distinguiamo tra carattere compensativo e carattere *ultra-compensativo* (“*reason-giving*”)<sup>70</sup>. Come abbiamo visto, riconosciamo carattere compensativo ai benefici esistenziali quando, data la scelta tra creare una persona e non crearla, consideriamo i dolori che questa patirà – danni esistenziali – alla stregua di ragioni per non crearla e consideriamo i piaceri di cui godrà – benefici esistenziali – come ragioni che solo neutralizzano – in questo senso compensano – le ragioni contrarie alla creazione. Se riconosciamo solo carattere compensativo, vi è un tetto alla rilevanza dei vari benefici esistenziali che una vita contiene, e questo tetto è rappresentato dall'ammontare dei danni esistenziali: una volta compensati questi danni, possiamo disinteressarci del saldo netto dei benefici – che una vita sia estatica, gioiosa, soddisfacente o appena degna di essere vissuta non ha rilevanza. Riconosciamo carattere *ultra-compensativo* ai benefici esistenziali, invece, qualora riconosciamo loro un valore che eccede la neutralizzazione dei danni esistenziali – per cui avrà rilevanza se una vita sia estatica o appena degna di essere vissuta.

Emerge però un problema. Se, nel considerare il beneficio esistenziale, gli attribuiamo solo valore compensativo, evitiamo la Conclusione Ripugnante – evitiamo anche che dovremmo “rimpiazzare” anziché “curare” –, ma perdiamo le basi per asserire che “agiatezza” sia meglio di “stenti”. Se, al contrario, attribuiamo valore *ultra-compensativo* al beneficio esistenziale, avremo le basi per preferire “agiatezza” a “stenti”, ma, apparentemente, andremo incontro alla Conclusione Ripugnante: ogni volta che la vita di una persona possibile conterrà una quantità di piaceri che eccede i dolori, riconosceremo un valore contributivo positivo alla creazione di quella vita e questo valore potrà compensare il decremento del benessere delle persone esistenti.

La domanda è allora se è possibile riconoscere carattere *ultra-compensativo* al beneficio esistenziale senza asserire anche che la creazione di vite aggiuntive con livello di benessere positivo abbia valore contributivo positivo. Distinguiamo due generi di casi

*Maggiore o minore.*

Prima scelta: Tizio esisterà e avrà benessere pari a 80, Caio non esisterà.

Seconda scelta: Tizio non esisterà, Caio esisterà e avrà benessere pari a 60.

*Qualcuno o nessuno.*

Prima scelta: Tizio esisterà e avrà benessere pari a 80.

Seconda scelta: Tizio non esisterà.

Un modo per dare valore alla dimensione *ultra-compensativa* del beneficio esistenziale è ritenere che, se una persona possibile avrà benessere positivo, allora abbiamo ragioni per creare quella persona – nel caso “qualcuno o nessuno” dobbiamo optare per la prima scelta. Se adottiamo questo principio è vero che incorriamo nella Conclusione Ripugnante. Tuttavia, possiamo dare valore alla dimensione *ultra-compensativa* del beneficio esistenziale anche ritenendo che, lungi dal costituire una ragione per creare una persona anziché *non crearla*, il benessere netto positivo contenuto nella vita di una persona possibile sia decisivo per stabilire se creare quella persona anziché *un'altra persona*. Chiamiamo questo principio *Comparativismo Condizionale*. L'idea di fondo è che non abbiamo ragione per creare persone felici, ma, se le creiamo – da qui “condizionale” – abbiamo ragioni per creare quelle che avranno una vita migliore: in “qualcuno o nessuno” le due alternative sono indifferenti; in “maggiore o minore” invece abbiamo ragioni per preferire la prima scelta<sup>71</sup>.

Per quanto promettente possa sembrare, il *Comparativismo Condizionale* implica conclusioni contro-intuitive. Prendiamo una versione rafforzata del Paradosso della Mera Addizione

<sup>70</sup> MCMAHAN 2009 53 ss.; 2013, 20 ss., 23 ss.

<sup>71</sup> La distinzione tra le scelte del tipo “maggiore o minore” e “qualcuno o nessuno” è presa da MCMAHAN 2013, 23. Il *Comparativismo Condizionale* si basa sull'idea che esistano beni condizionali. Su beni condizionali e beni *categorici* WILLIAMS 1973a e il relativo commento in BROOME 2004, 152-157.

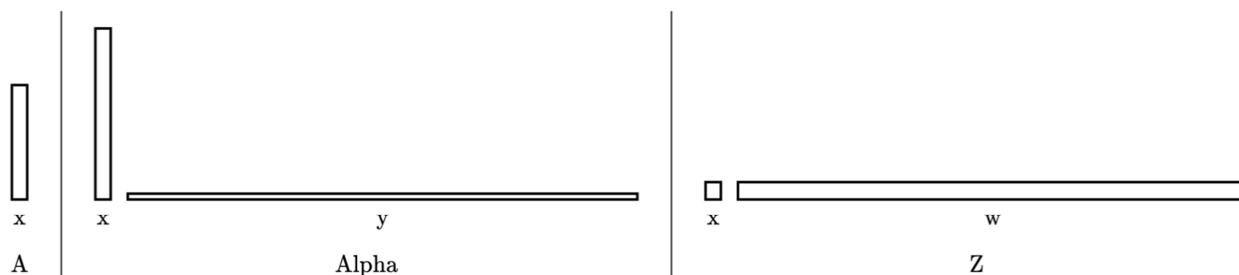


Fig.7.

Il mondo Alpha costituisce una variante di A+. In Alpha il benessere dei più fortunati non è eguale ma addirittura molto superiore a quello delle persone di A. Si tratta non di una “mera addizione” ma di una “addizione benevola”<sup>72</sup>. Nell’esempio che ci interessa vi è un dettaglio aggiuntivo. I grafici questa volta ci indicano quali persone – nome e cognome – esisteranno nei vari scenari. Questa indicazione ci viene fornita dalle lettere  $x, y, w$ . Così sappiamo che le stesse persone del gruppo  $x$  esisteranno in tutti gli scenari, mentre le persone del gruppo  $y$  solo nello scenario Alpha e le persone del gruppo  $w$  solo nello scenario Z.

Sulla base del Comparativismo Condizionale, se parliamo A e Alpha dovremmo concludere che Alpha sia meglio di A. Infatti, in Alpha non vengono prodotte vite con benessere netto negativo –vincolo che ci portiamo dietro sin dal Comparativismo Asimmetrico – e vi è una notevole quantità di benefici comparativi a vantaggio del gruppo  $x$ . Se, poi, parliamo Alpha e Z, sembrerebbe doversi concludere che Z sia migliore di Alpha. Vero è che le persone che compongono il gruppo  $x$  patiscono un danno comparativo, ma è anche vero che – qualora si producesse Z – le persone che compongono il gruppo  $w$  ottengono un vantaggio non-comparativo maggiore di quello che ottengono le persone del gruppo  $y$  – qualora si producesse Alpha –, e, se riconosciamo la sussistenza di benefici esistenziali ultra-compensativi nei casi “maggiore o minore”, allora dobbiamo attribuire valore a questa circostanza. Quanto valore?<sup>73</sup> Senza addentrarci in una simile questione, basta dire che se anche gli attribuiamo un valore commensurabile con quello dei danni comparativi, allora esisterà sempre un numero di persone il cui beneficio non-comparativo aggregato superi il danno comparativo patito da un minor gruppo di persone. Se, infine, parliamo A e Z, A è migliore di Z – infatti il beneficio non-comparativo nel Comparativismo Condizionale non ha valore ultra-compensativo nei casi del tipo “qualcuno o nessuno” e l’alternativa tra A e Z è appunto un caso del tipo “qualcuno o nessuno”.

Il Comparativismo Condizionale non implica la Conclusione Ripugnante, ma genera – in questo come in altri casi<sup>74</sup> – relazioni intransitive: sebbene Alpha sia migliore di A, e Z sia migliore di Alpha, Z è peggiore di A.

Per chiunque ritenga che il Principio di Transitività sia una condizione di accettabilità di qualsiasi soluzione, principio o teoria in ambito pratico, il Comparativismo Condizionale si rivela inadeguato<sup>75</sup>.

<sup>72</sup> L’espressione “addizione benevola” può ritrovarsi in HUEMER 2008, 901 ss.

<sup>73</sup> Il problema è discusso in MCMAHAN 2013, 15 ss.

<sup>74</sup> Diverso è, ad esempio, il caso adoperato da McMahan contro il Comparativismo Condizionale, MCMAHAN 2013, 25 ss.

<sup>75</sup> Questa critica generale al Comparativismo – escluso quello che abbiamo chiamato Comparativismo Morbido – è stata svolta da BROOME (2004, 141 ss., in particolare 146-149); per una critica al Comparativismo Condizionale, *ivi* 152-157, in particolare 156. Per una disamina accurata TEMKIN 2012, 417 ss. Temkin utilizza un esempio diverso da quello delle Addizioni Benevole, tuttavia alle addizioni benevole fa espressamente riferimento a p. 428. Inoltre, in quel contesto, Temkin non analizza e difende il Comparativismo Condizionale esattamente così come da me configurato sulla scorta di McMahan, ma una versione simile TEMKIN 2012, 437 ss.

## 7. Intransitività, impossibilità

Secondo Larry Temkin, in certi casi – tra cui il Paradosso della Mera Addizione –, attribuiamo a certi valori morali carattere *essenzialmente comparativo*, con ciò intendendosi che, per stabilire come una data popolazione si classifica rispetto a questi valori, non è sufficiente osservare le caratteristiche intrinseche della popolazione medesima<sup>76</sup>. Viceversa, la realizzazione di questi valori rispetto a una certa popolazione dipenderà dalle popolazioni alternative a cui viene comparata – la realizzazione di quel valore in una data popolazione sarà sempre realizzazione di quel valore in una data popolazione per come comparata a un'altra popolazione o a una pluralità di altre popolazioni<sup>77</sup>.

Vediamo di precisare. Ogni popolazione presenta un insieme di *caratteristiche* o *fattori* – l'esservi persone che stanno peggio di altre, l'esservi una certa quantità di benessere, ecc. Ogni popolazione, poi, può essere giudicata sulla base di diversi *valori* – Principio Totale, Maximin, valore dell'Eguaglianza, ecc. Secondo un modo di vedere diffuso, come una data popolazione si classifichi rispetto a un certo valore dipende solo dalle caratteristiche di questa – i vari fattori attribuiscono valore sempre allo stesso modo, la loro rilevanza è costante. Diciamo allora che i valori sono *intrinseci*. Invece, riconoscere che certi valori siano essenzialmente comparativi significa riconoscere che come un certo valore si realizzi in una data popolazione non dipende solo dalle caratteristiche di questa, ma anche dalle caratteristiche delle popolazioni a cui la prima è comparata – i vari fattori non rileveranno allo stesso modo, ma assumeranno o perderanno rilevanza a seconda delle alternative considerate<sup>78</sup>.

Per illustrare il punto torniamo all'esempio del Maximin – per cui la popolazione comparata migliore è quella in cui i più svantaggiati stanno meglio (§2).

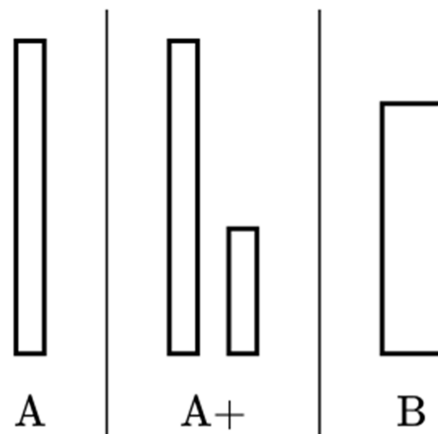


Fig. 8

I più svantaggiati di A – ossia tutti i membri di A – stanno meglio dei più svantaggiati di B – anche qua tutti i membri di B. Dato questo fattore, A è meglio di B rispetto al Maximin. Per lo stesso fattore, B è meglio di A+ rispetto al Maximin. Infatti, i più svantaggiati di A+ stanno peggio dei più svantaggiati di B. I più svantaggiati di A+ stanno peggio dei più svantaggiati di A. Cosa dire di quest'ultima circostanza? Se ritenessimo che i valori sono sempre intrinseci – dipendono

<sup>76</sup> Più esattamente «*The Essentially Comparative View of Moral Ideals (ECVMI)*: Roughly, there is at least one moral ideal, *F*, and at least one outcome, *O*, such that there is no answer to the question of how good *O* is regarding *F*, based solely on *F*'s internal features; or, even if, for each moral ideal, *F*, and each outcome, *O*, there is an answer to the question of how good *O* is regarding *F* based solely on *O*'s internal features, there is at least one ideal, *F*, and two outcomes *O*<sub>1</sub> and *O*<sub>2</sub>, such that how *O*<sub>1</sub> compares with *O*<sub>2</sub> regarding *F* is not simply a function of the extent to which *O*<sub>1</sub> is good with respect to *F* based solely on *O*<sub>1</sub>'s internal features and the extent to which *O*<sub>2</sub> is good with respect to *F* based solely on *O*<sub>2</sub>'s internal features» (TEMKIN 2012, 371).

<sup>77</sup> TEMKIN 2012, 368 ss.

<sup>78</sup> TEMKIN 2012, 228.



sempre dagli stessi fattori interni dei vari scenari – allora diremmo anche che A+ sia peggio di A rispetto al Maximin. Tuttavia, questo approccio trascura il fatto che A+ sia una mera addizione rispetto ad A, una circostanza a cui molti di noi – si vuole sia un dato di fatto<sup>79</sup> – sono sensibili. Infatti, l’attaccamento di molti di noi verso il Maximin si traduce nella volontà di migliorare le vite di coloro che sono svantaggiati, ma non anche nell’impegno di astenerci dal mettere al mondo persone che hanno vite degne di essere vissute. Sebbene in A+ esistano persone che stanno peggio sia di tutti coloro che popolano A sia di tutti coloro che popolano B, se questa circostanza incarna un disvalore dal punto di vista del Maximin dipenderà dal se compariamo A+ con B, o A+ con A. Se compariamo A+ con B questa circostanza avrà valore, ma se compariamo A+ con A questa circostanza non avrà valore: A+ sarà peggiore per il Maximin di B, ma non anche di A<sup>80</sup>. In questo senso attribuiamo al Maximin carattere essenzialmente comparativo.

Questa impostazione è gravida di conseguenze. Secondo il *Principio di Transitività e Sostituzione* dati tre scenari x,y,z, se x è meglio di y rispetto a un certo valore, e y è meglio di z rispetto a quello stesso valore, allora x sarà meglio di z rispetto a quel valore; se poi x è equivalente a y rispetto a un certo valore, allora, qualsiasi sia la relazione tra x e z rispetto a quel valore, y e z avranno quella stessa relazione<sup>81</sup>. Se per comparare assiologicamente x,y,z rispetto a un certo valore rilevano i medesimi fattori o caratteristiche, allora ci si può aspettare che la relazione sia transitiva; ma se, quando compariamo x e y, rispetto a un certo valore, rileva un insieme di fattori *diverso* da quello che rileva quando compariamo, rispetto al medesimo valore, y e z, allora la relazione non sarà più transitiva – non potremo inferire, dalle precedenti, la relazione tra x e z – e questo anche se la relazione assiologica non sia imprecisa<sup>82</sup>.

Così, tornando al caso del Maximin, sebbene, rispetto al Maximin, A sia migliore di B e B sia migliore di A+, A non sarà migliore di A+, in deroga al Principio di Transitività e Sostituzione. Le relazioni assiologiche basate su un valore essenzialmente comparativo non saranno relazioni transitive (e sostituibili)<sup>83</sup>.

Dove ci portano queste considerazioni sul carattere essenzialmente comparativo dei valori? Secondo il Paradosso della Mera Addizione tre nostre intuizioni profondamente radicate – il Principio della Mera Addizione, il Principio di Non-Antieguagliarismo e il Rigetto della Conclusione Ripugnante – sono contraddittorie (§2). Questa incoerenza tuttavia presuppone l’accettazione del Principio di Transitività e Sostituzione. Se il Principio di Transitività e Sostituzione non trova applicazione potremo concludere che A+ sia almeno tanto buono quanto A, e che Z sia migliore di A+, senza dover inferire che Z sia migliore di A<sup>84</sup>.

<sup>79</sup> «This merely expresses what we really are», TEMKIN 2012, 381, 407.

<sup>80</sup> TEMKIN 2012, 402-416.

<sup>81</sup> TEMKIN 1987, 143 s. «Given any concept c, for all x, y, and z to which c is appropriately applied, if x is c-better than y, and y is c-better than z, then x is c better than z. [...] Given any concept c, for all x, y, and z to which c is appropriately applied, if x is as c-good as, or c-equivalent to y, then however x c compares to z, that is how y c compares to z». V. anche TEMKIN 2012, 378.

<sup>82</sup> TEMKIN 2012, 378 s. Si ricorderà (§5) che l’imprecisione – Temkin parla di “vaghezza” – introduce la possibilità che vi siano relazioni assiologiche intransitive. Uno degli aspetti più suggestivi del meticoloso lavoro di Temkin consiste nel mostrare come anche relazioni precise possano derogare al Principio di Transitività.

<sup>83</sup> Per il meccanismo generale delle deroghe al Principio di Transitività e Sostituzione TEMKIN 2012, 171-175, 181, rispetto ai valori morali *ivi*, 369,405,410,416. Oltre a violare il principio di Transitività e Sostituzione, i valori essenzialmente comparativi violeranno anche il *Principio dell’Indipendenza delle Alternative Irrilevanti*, per cui se due alternative stanno in una certa relazione – ad esempio, la prima alternativa è migliore della seconda – rispetto a un dato valore quando sono le uniche due alternative, questa relazione non cambierà se anche altre alternative sono disponibili. Per esemplificare una violazione del Principio dell’Indipendenza delle Alternative Irrilevanti si cita spesso una battuta di Morgenbesser “vorrei il gelato alla fragola, salvo abbiate quello alla vaniglia, in tal caso prendo quello al cioccolato” (TEMKIN 2012, 367-390).

<sup>84</sup> TEMKIN 2012, 394.

Il senso che Temkin ha voluto dare a questo ragionamento non è tuttavia necessariamente confortante. Di alcune relazioni ci aspettiamo precisamente che siano non-transitive: se Tizio è innamorato di Caio e Caio è innamorato di Sempronio, non segue che Tizio sia innamorato di Sempronio. Cionondimeno, ci aspettiamo che certe relazioni assiologiche – l'essere uno scenario migliore di un altro rispetto al Maximin, oppure rispetto al Principio di Pareto, o rispetto al valore dell'Eguaglianza, ecc. – siano transitive. Mettere a fuoco il carattere essenzialmente comparativo di certi valori significa tradire la nostra aspettativa in favore della transitività. Poiché, poi, queste relazioni assiologiche particolari – rispetto al Maximin, rispetto al Principio di Pareto, rispetto al valore dell'Eguaglianza, ecc. –, che non sarebbero transitive, sono i fattori (fattori di livello superiore a quelli da cui le medesime dipendono a loro volta) da cui dipende la relazione “x è meglio di y *tutto considerato*” – relazione assiologica ultimativa –, segue che anche questa relazione, che ci aspetteremmo sia transitiva, erediti la non-transitività delle prime<sup>85</sup>.

Il Paradosso della Mera Addizione, allora, sarebbe il contesto di scoperta di un paradosso ben più grande: siamo catturati da due visioni incoerenti dei valori. Per la prima alcuni valori morali sono essenzialmente comparativi – nel comparare, rispetto a certi valori, alcuni scenari, non sempre diamo rilevanza agli stessi fattori–; per la seconda tutti i valori morali sono intrinseci – ci aspettiamo che quando uno scenario sia valutato migliore di un altro rispetto a certi valori o in senso ultimativo, tutto considerato, si instauri una relazione transitiva<sup>86</sup>.

Mettere in discussione il Principio di Transitività non conduce dunque alla risoluzione dei paradossi dell'etica della popolazione, quanto alla constatazione di incoerenze del giudizio più vaste di quelle inizialmente mostrate dal Paradosso della Mera Addizione. Similmente Gustaf Arrhenius ha mostrato in modo formale – attraverso diversi Teoremi dell'Impossibilità – come, anche trovando dei principi che rispettino grosso modo il Principio della Mera Addizione, il Principio di Non-Antieguaglianza e il Rigetto della Conclusione Ripugnante, questi principi si scontrino con altre nostre intuizioni altamente plausibili<sup>87</sup>.

<sup>85</sup> Mentre di certe relazioni non ci aspettiamo siano transitive – sopra ho fatto un esempio con la relazione “innamorato di” –, vi sono altre relazioni che ci aspettiamo siano transitive TEMKIN 2012, 433, 473 s., 487. Nell'opera di Temkin questa aspettativa di transitività non è estesamente trattata. Una delle ragioni, prese in considerazione da Temkin, per cui ci aspettiamo che la transitività non patisca deroghe nel ragionamento pratico deriva dalla credenza che preferenze intransitive generino il problema, noto agli economisti, della “*money pump*”. A tal proposito Temkin – facendo tesoro dell'idea di *pre-commitment* (ELSTER 1976) – ritiene che il problema della *money pump* sia più teorico che pratico, TEMKIN 2012, 183-193. Sulla circostanza per cui la non transitività di una relazione viene ereditata dalla relazione assiologica ultimativa si veda pp. 203-214. Si noti anche come proprio il valore dell'Eguaglianza secondo Temkin non sia essenzialmente comparativo, sebbene sia utile trattarlo come tale a scopi illustrativi (376).

<sup>86</sup> TEMKIN 2012, 390-400. È importante sottolineare come il riconoscimento del carattere essenzialmente comparativo di molti valori morali non sia un artificio teorico volto a risolvere il Paradosso della Mera Addizione. Come si è detto prima, tale riconoscimento è semplicemente la presa d'atto del nostro modo di attribuire valore a certi fattori - «This merely expresses what we really are». Qualora ritenessimo che tutti i valori sono intrinseci, potremmo salvare il Principio di Transitività; di più, i paradossi dell'etica della popolazione risulterebbero meno inquietanti. Per fare un esempio, se ritenessimo che tutti i valori sono intrinseci e dovessimo comparare A, A+ e B (oppure A, A+ e Z) dovremmo svolgere considerazioni del genere: A+ è migliore di A per quanto riguarda il Principio Totale, ma è peggiore per quanto riguarda il valore dell'Eguaglianza e il Maximin. (Infatti, per ritenere che A+ non sia peggiore di A rispetto a Eguaglianza e Maximin dobbiamo ricordarci che A+ è una *mera addizione* rispetto ad A, ma ora stiamo provando a ragionare come se tutti i valori fossero intrinseci e questo aspetto è irrilevante) Sarà A+ migliore o peggiore di A? dipende dal *peso* che riterremo si debba attribuire ai vari valori in conflitto. Lo stesso procedimento poi andrà applicato alle relazioni tra A+ e B (o tra A+ e Z). Certo, non è affatto garantito che le relazioni assiologiche ultimative tra le popolazioni alternative siano facilmente determinabili, ma il problema sarebbe diventato un problema tradizionale - avremmo una pluralità di valori in conflitto l'uno con l'altro che ingenerano relazioni assiologiche vaghe, non più un insieme di relazioni assiologiche ciascuna in sé intuitiva ma complessivamente contraddittorie (come accennato in §2). Dunque, come si diceva, se ritenessimo che tutti i valori sono intrinseci, salveremmo il Principio di Transitività e Sostituzione rispetto a relazioni per le quali ci aspettiamo che valga e ridurremmo l'intrattabilità dei paradossi dell'etica della popolazione, però, appunto, faremmo un torto al modo in cui accade che diamo valore ai fatti del mondo.

<sup>87</sup> Alcune delle confutazioni di Arrhenius dei principi avanzati sono state esposte nei precedenti paragrafi. Ripeter-

## 8. *Accettare la Conclusione Ripugnante*

Secondo alcuni autori il Paradosso della Mera Addizione e in generale i paradossi dell'etica della popolazione andrebbero risolti accettando la Conclusione Ripugnante. In questo paragrafo verranno descritti alcuni degli argomenti che sono stati avanzati in questa direzione. Invece, nel prossimo paragrafo si metterà a confronto quest'ultima famiglia di argomenti con gli argomenti precedentemente analizzati, e si formuleranno alcuni problemi che non sono stati precedentemente discussi nel dibattito sull'etica della popolazione.

Il primo argomento è stato sviluppato da Torbjörn Tännsjö. Quando compariamo il mondo A – un mondo in cui ciascuno ha un livello di benessere molto alto – e il mondo Z – un mondo molto più popolato del primo in cui ciascuno vive una vita che è appena degna di essere vissuta –, siamo portati a pensare che le nostre vite – esattamente, le vite dei membri privilegiati delle nazioni più ricche del pianeta – siano in qualche modo simili a quelle che si vivrebbero in A e certamente ben migliori di quelle di coloro che vivono in Z. Per Tännsjö nel pensare questo ci inganniamo. L'idea di fondo è che, una volta garantiti i bisogni primari, una vita diventa degna di essere vissuta e qualsiasi incremento ulteriore della ricchezza materiale ne muove di poco il livello di benessere reale. Alcuni fatti – si precisa – sono effettivamente in grado di incrementare il benessere delle persone, ma questi fatti non hanno nulla a che fare con la ricchezza materiale e sono per loro natura effimeri – l'infatuazione sentimentale ricambiata, l'appropriata somministrazione di droghe. Dato il loro carattere effimero questi fatti non determineranno grandi differenze nei livelli di benessere delle persone. Vite riconoscibilmente umane con un alto livello di benessere positivo non sono immaginabili. Tirando le somme, molte persone oggi giorno non vedono loro garantiti nemmeno i bisogni primari, ma queste persone hanno vite degne di *non* essere vissute – non è a loro che dobbiamo guardare per immaginare Z –; coloro ai quali i bisogni primari sono garantiti vivono il più delle volte vite appena degne di essere vissute – io e voi stiamo tutti “schiacciati” attorno al livello Z<sup>88</sup>. Come non vi è nulla di ripugnante nelle migliori condizioni di vita riconoscibilmente umane, così non vi è nulla di ripugnante in Z, e allora il Paradosso della Mera Addizione va inteso, invece che come un paradosso, come un argomento a favore dell'accettazione di quella che solo impropriamente viene chiamata “Conclusione Ripugnante” – in questo senso, dobbiamo accettare la Conclusione Ripugnante<sup>89</sup>.

A questo argomento potrebbe risponderci che la maggior parte di noi può immaginare eventi drammatici – per esempio, la perdita di un arto – che peggiorano la qualità della nostra vita senza pregiudicare il fatto che questa resti degna di essere vissuta. Eppure, ciò sarebbe inspiegabile laddove in generale vivessimo vite appena degne di essere vissute – come sarebbe possibile che le nostre vite appena degne di essere vissute possano patire evidenti peggioramenti e pur preservarsi vite degne di essere vissute? La risposta a questa obiezione è – a mio modo di vedere – significativa: spesso le persone che perdono qualcosa che credevano indispensabile per condurre un'esistenza decente, dopo una fase di iniziale smarrimento, si adattano alla loro nuova condizione e tornano al livello di benessere che avevano prima. Insomma, anche quando crediamo che, perdendo molte delle cose che abbiamo, il nostro benessere diminuirà durevolmente, noi crediamo il falso – ci stiamo illudendo.

Il secondo argomento sul quale voglio soffermarmi è stato presentato da Michael Huemer. Si tratta di un esperimento mentale diviso in due fasi. Prima fase. Immaginiamo una coppia che ha già due figli e si stia interrogando sull'opportunità di farne un terzo. Verosimilmente la vita del terzo figlio avrà benessere positivo, tuttavia il benessere della coppia patirà un leggero decremen-

correre anche solo uno dei suoi teoremi dell'impossibilità richiederebbe troppo spazio. Per un esempio risalente ma comunque esemplificativo dell'operazione teorica di Arrhenius si rinvia ad ARRHENIUS 2000a e ARRHENIUS 2000b.

<sup>88</sup> TÄNNSJÖ 2004, 223-226.

<sup>89</sup> TÄNNSJÖ 2004, 232-235.

to. Un amico, interpellato sul punto, sentenzia che sarebbe bene che la coppia non abbia il terzo figlio. Le ragioni della risposta dell'amico sono abbastanza comprensibili: il concepimento del terzo figlio sarà un *male* per la coppia, e *non* sarà un *bene* per il terzo potenziale figlio – in linea con l'idea per cui i benefici esistenziali non riguardino le scelte “qualcuno o nessuno” (§6, in particolare il Comparativismo Condizionale<sup>90</sup>). Seconda fase dell'esperimento mentale. La coppia ha avuto una bambina, si chiama Sally. Sally ha una vita ben degna di essere vissuta, la coppia sta un po' peggio di prima. Tutti, i genitori, i tre figli e l'amico, stanno pranzando insieme. L'amico, che non ha avuto ragione di ricredersi, sentenzia che sarebbe stato bene che Sally non fosse nata.

Verosimilmente, mentre il primo giudizio verrebbe accolto come ragionevole, il secondo verrebbe accolto come irragionevole e offensivo. Eppure le cose sono andate proprio nel modo in cui ci si aspettava che andassero quando il primo giudizio era stato emesso. Secondo Huemer, saremmo vittime di un *pregiudizio presentista*: siamo propensi a riconoscere il valore della creazione di una persona la cui vita è degna di essere vissuta, una volta che la persona sia stata creata – atteggiamento retrospettivo; ma, quando si tratta di creare una persona, non siamo particolarmente impressionati dall'immaginare che questa avrà una vita degna di essere vissuta – atteggiamento prospettivo. Ma se riconosciamo il valore retrospettivamente allora dobbiamo riconoscerlo anche prospettivamente: abbiamo ragioni per creare nuove vite con benessere positivo anche se questo si traduce in un decremento nel benessere delle persone esistenti. Per queste ragioni dobbiamo accettare la Conclusione Ripugnante<sup>91</sup>.

### 9. Note su alcuni problemi di metodo nel ragionamento morale

Il tratto comune di questi argomenti in favore dell'accettazione della Conclusione Ripugnante è l'attacco al nostro punto di vista, o, forse meglio, a una “parte” del nostro punto di vista: siamo in preda a illusioni e di queste illusioni non possiamo sbarazzarci. Con riferimento al primo argomento, ascoltare la testimonianza di chi, dopo un grave incidente, racconta di come alla fine gli sembri di essersi adattato alle sue nuove condizioni di vita, non basta per farci cessare di concepire i gravi incidenti come eventi che diminuiscono il benessere delle loro vittime; con riferimento al secondo argomento, constatare come si tenda ad attribuire valore in sé alla vita di persone esistenti non basta per farci attribuire quello stesso valore alle persone meramente possibili che potremmo creare. Il concepirci come persone che potrebbero perdere molto benessere e pur continuare a vivere vite degne di essere vissute, il non essere empatici nei confronti delle persone possibili che potrebbero esistere e invece non esisteranno, se sono illusioni, sono illusioni *persistenti* – illusioni che non scompaiono una volta che queste nostre condizioni mentali e i sottostanti meccanismi psicologici ci sono stati messi davanti.

La tesi che intendo avanzare è che questa mossa – il richiamo alle influenze distorsive del giudizio, illusioni persistenti – sembra, almeno a prima vista, eccentrica rispetto a un certo modo di intendere il ragionamento morale. Per capire in che modo sia eccentrica facciamo un passo indietro.

In etica della popolazione – e lo stesso vale per altri dibattiti etico-sostanziali – ricorrono espressioni come intuizione, intuitivo, contro-intuitivo – si dice, a esempio, che questo o quel giudizio è “intuitivo” o “contro-intuitivo”, che questo o quel principio ha conseguenze, applicazioni, “intuitive” o “contro-intuitive”. In metaetica esiste un uso del termine “intuizione morale” per cui l'intuizione morale è la facoltà attraverso la quale possiamo accedere alla conoscenza

<sup>90</sup> O magari sarà un bene per il terzo potenziale figlio, ma il beneficio esistenziale avrebbe un valore lessicalmente inferiore al beneficio comparativo di cui si avvantaggerebbe la coppia qualora si astenesse dal fare il terzo figlio. Questa è forse l'ipotesi considerata da Huemer, tuttavia possiamo tralasciare queste complicazioni.

<sup>91</sup> HUEMER 2008, 916-918. Huemer offre ulteriori argomenti in favore dell'accettazione della Conclusione Ripugnante.

delle verità morali o – considero le due espressioni come sinonime – dei valori morali oggettivi. L’Intuizionismo, in questo senso, è una posizione metaetica oggettivista<sup>92</sup>. A seconda di come questa concezione di intuizione venga sviluppata, l’essere un giudizio intuitivo potrà essere condizione necessaria e sufficiente perché il giudizio rifletta una verità morale, o potrà essere condizione solo necessaria, o solo sufficiente, o magari i giudizi intuitivi saranno solo il punto di partenza per arrivare, attraverso procedimenti non intuitivi, alla conoscenza di certe verità morali. A seconda di come venga sviluppata, potrà ammettersi che vi siano intuizioni corrette e intuizioni scorrette. Ma, come si voglia modulare l’Intuizionismo, in ogni caso il presupposto resta che vi sono dei valori morali oggettivi e le intuizioni sono uno strumento per la loro conoscenza.

Questa prima concezione di intuizione – con il suo legame indissolubile con l’idea di oggettivismo morale – deve essere considerata solo per poterla mettere da parte. La seconda concezione, quella che ci interessa, prescinde da una posizione sull’oggettività dei valori morali. Cosa resta una volta eliminato il legame con l’oggettività dei valori morali? Resta che l’intuizione morale è una risposta emotiva di approvazione o di disapprovazione rispetto a una certa condotta o stato di fatto. Una risposta emotiva rapida? Automatica? Inconscia? Non-sillogistica? Necessariamente motivante? Per lo scopo che mi prefiguro non credo sia necessario qualificare ulteriormente cosa sia un’intuizione se non dicendo che quando abbiamo un’intuizione a favore di un certo giudizio morale – quando quel giudizio ci sembra intuitivo – quel giudizio ci sembra corretto. E viceversa, quando abbiamo un’intuizione contro un certo giudizio morale – quando quel giudizio ci sembra contro-intuitivo – quel giudizio ci sembra scorretto. Una caratteristica di questo uso del termine intuitivo è che l’agente che ha un’intuizione a favore di un certo giudizio può criticare questa sua risposta – l’intuizione – e ritenere che il giudizio sia scorretto o, almeno, che sia un giudizio di cui valga la pena dubitare. Infatti, comunemente a molti di noi capita di ritenere che qualcosa che ci *sembra* corretto, *effettivamente* non lo sia – fermo restando che (come per l’idea di “intuizione”) sto assumendo un’idea di “correttezza” e, dunque, di “effettivamente corretto” sufficientemente indeterminata da risultare compatibile anche con posizioni che negano l’oggettività dei valori morali<sup>93</sup>.

Tenendo a mente questa seconda nozione di intuitività, possiamo delineare un primo metodo del ragionamento morale secondo il quale una soluzione intuitiva è una soluzione corretta. Secondo questo metodo, tra intuitività e correttezza non c’è iato: l’essere un giudizio intuitivo è condizione necessaria, ma anche sufficiente, perché quel giudizio sia corretto. Giudizi contro-intuitivi non potranno essere corretti: almeno per quanto riguarda la scelta, o la contemplazione, dei nostri valori supremi (*telici*), non esistono pregiudizi. Chiamiamo questo primo metodo del ragionamento morale “Metodo Intuizionista”<sup>94</sup>.

A eccezione di quelli in favore dell’accettazione della Conclusione Ripugnante, gli argomenti e le obiezioni che sono stati passati in rassegna (§3-7) sono sussumibili entro il Metodo Intuizio-

<sup>92</sup> HUEMER 2005, 6 ss.

<sup>93</sup> In che misura un agente che nega l’oggettività dei valori morali può ritenere che un certo giudizio sia – o non sia – effettivamente corretto? Non può essere questa l’occasione per esaminare questo problema con l’attenzione che meriterebbe. Tuttavia, ritengo che un agente che neghi l’oggettività dei valori morali possa dire che un certo giudizio sia effettivamente corretto *per lui* nel momento in cui attiva dei processi di controllo su quel giudizio e questi processi lo riconfermano, e poi attiva processi di controllo sui processi di controllo e questi li riconfermano, e così via ricorsivamente, fino all’esaurimento delle sue risorse psicologiche (la sua concentrazione prima di tutto). Su giudizi normativi e ricorsività dei processi di controllo, BRIGAGLIA 2018.

<sup>94</sup> Per valore *telico* intendo un valore finale contrapposto a un valore strumentale: dormire ha un valore strumentale al piacere, il piacere è invece un valore finale. Questo uso del termine “telico” è tratto dalle recenti opere di teoria morale generale di Parfit, Per esempio, «Our desires are *teleological* or *telic* when we want some event as an *end*» e «Such reasons are *telic* when they are provided by the facts that make some possible event good as an *end*» (PARFIT 2011, 44, 52). Secondo il Metodo Intuizionista non esistono pregiudizi sui valori telici, ma possono esistere pregiudizi sui valori strumentali: un agente morale può ingannarsi su quali siano i mezzi migliori per raggiungere il proprio scopo.

nista: i principi che costituiscono il contenuto dei vari argomenti, per avere successo, devono avere implicazioni che non siano contro-intuitive; mentre le obiezioni, per avere successo, devono mostrare che in realtà quei principi hanno implicazioni contro-intuitive<sup>95</sup>.

Davanti a questa ricostruzione potrebbe sorgere un dubbio. Se è vero che il Paradosso della Mera Addizione mostra come tre giudizi intuitivi sono tra di loro contraddittori, quali margini, in radice, vi potrebbero mai essere per il Metodo Intuizionista di giungere a una soluzione del paradosso? In fondo – si potrebbe dire –, se, da un lato, il Paradosso della Mera Addizione mostra come ci sembrano corrette tre posizioni contraddittorie e, dall'altro lato, il Metodo Intuizionista stabilisce che una soluzione che sembra corretta è una soluzione effettivamente corretta, non giungiamo forse all'*impasse* per cui abbiamo un insieme di giudizi corretti ma contraddittori? Direi che vi sono diversi modi in cui potremmo sperare di poter risolvere il Paradosso della Mera Addizione coerentemente con il Metodo Intuizionista. Una strategia ricorrente è, ad esempio, quella di provare a riformulare uno dei principi intuitivi che costituiscono le premesse del Paradosso della Mera Addizione in modo che il principio riformulato in sé resti intuitivo ma non risulti contraddittorio con gli altri principi. Ad esempio, il modo in cui PVV (§3) deroga – in questo senso riformula – al Principio della Mera Addizione, può sembrare, di per sé, intuitivo; e il principio riformulato non è in contraddizione né con il Principio di Non-Antieguaglianza, né con il Rigetto della Conclusione Ripugnante. Sebbene siano concepibili margini di speranza per una soluzione, la storia ha mostrato come i vari principi proposti – in sé almeno non platealmente contro-intuitivi e tra loro componibili – ingenerino altre conclusioni altamente contro-intuitive. Ad esempio, PVV ingenera la Conclusione Sadica. Abbiamo dunque una serie di argomenti che sembrano assumere come regole del gioco che un principio, o una teoria, ha successo se non ha implicazioni contro-intuitive, merita di essere scartato qualora ha implicazioni contro-intuitive – ciò è compatibile con il Metodo Intuizionista.

Gli argomenti in favore dell'accettazione della Conclusione Ripugnante che sono stati brevemente passati in rassegna nel paragrafo precedente, invece, non sembrano – non a prima vista almeno – compatibili con quello che ho chiamato Metodo Intuizionista. La premessa metodologica sulla quale riposano è, infatti, che alcune delle nostre intuizioni siano dei meri pregiudizi – illusioni persistenti – e che come tali, lungi dal costituire il banco di prova delle varie soluzioni prospettabili, andrebbero neutralizzate affinché non pervertano il ragionamento morale. Non tutti i nostri giudizi intuitivi sono corretti e, per converso, alcuni giudizi contro-intuitivi sono corretti e vanno accettati. Chiamiamo “Metodo dell'Accettazione” quel metodo del ragionamento per cui l'essere un giudizio intuitivo – ossia il sembrare quel giudizio corretto – e l'essere un giudizio giusto – l'essere quel giudizio effettivamente corretto – non coincidono e può darsi il caso che si debba accettare un giudizio contro-intuitivo.

La sussistenza di due diversi metodi del ragionamento morale crea un cortocircuito: il dibattito tende a diventare una discussione tra sordi. Prendiamo un utilitarista – come Tännsjö<sup>96</sup> – sostenitore del Principio Totale e seguace del Metodo dell'Accettazione. L'utilitarista riterrà non solo che si debba accettare la Conclusione Ripugnante, ma anche che per stabilire il valore di una vita si debba solo fare riferimento alla somma algebrica dei piaceri e dei dolori che questa contiene (§1). Un critico, seguace del Metodo Intuizionista, può prontamente mostrare a questo utilitarista come la sua posizione vada incontro a conclusioni ancora

<sup>95</sup> Gli argomenti e le obiezioni passati in rassegna nei paragrafi 3-7 non sono solo compatibili con il Metodo Intuizionista. Certamente sarebbero compatibili pure con una variante del Metodo Intuizionista per cui l'essere un certo giudizio intuitivo è condizione necessaria ma non anche, o non sempre, sufficiente per rendere quel giudizio giustificato. Questa variante – al pari della versione del Metodo Intuizionista da me considerata – escluderebbe che un giudizio contro-intuitivo possa essere corretto e questo basterebbe al rendere compatibile il metodo con gli argomenti passati in rassegna nei paragrafi 3-7.

<sup>96</sup> TÄNNSJÖ 1998.

più contro-intuitive della Conclusione Ripugnante. Implicherebbe, ad esempio, che un mondo in cui miliardi di persone vivono vite lunghe e gioiose sia peggiore di un mondo in cui esiste un numero molto maggiore di creature che sperimentano solo gioia ma vivono ciascuna per appena un minuto. Avremmo ragioni per farci da parte e mettere al posto nostro un numero imprecisato di queste creature? Per molti si tratta di una conclusione decisamente contro-intuitiva – una Conclusione Ripugnante Rafforzata. Eppure un utilitarista seguace del Metodo dell'Accettazione può rispondere che l'errore del suo interlocutore è sempre lo stesso, quello di dare troppo credito alle intuizioni: quel che la Conclusione Ripugnante Rafforzata mostra è solo quanto le nostre intuizioni potrebbero essere irrimediabilmente compromesse. Non è forse a causa di un pregiudizio specista – si dirà – che non riusciamo a vedere il maggior valore contenuto in questo mondo di mosche felici?<sup>97</sup> Mi sembra che i binari comunicativi siano ormai diventati paralleli.

Alcune notazioni conclusive problematiche. Non si ha qui l'intento di difendere il Metodo Intuizionista, tuttavia, almeno apparentemente, le regole del Metodo Intuizionista sono abbastanza determinate. Sono regole semplicissime: se la soluzione prospettata è intuitiva, la soluzione è corretta; finché non raggiungiamo una situazione in cui tutti i giudizi che troviamo intuitivi possono coesistere armonicamente non abbiamo finito il nostro lavoro; potrebbe darsi che sia impossibile trovare una soluzione accettabile dal momento che, considerate le varie combinazioni disponibili di giudizi coerenti, non ve ne è, tra queste, una in cui tutti i giudizi siano da noi trovati intuitivi<sup>98</sup>. Per il Metodo dell'Accettazione il discorso è diverso. Quali sono le regole che sovrintendono alla identificazione dei pregiudizi è una questione che non è chiara e che deve essere chiarita. Perché dovremmo distaccarci da ciò che ci sembra corretto in una direzione e non nella direzione opposta? In mancanza di una risposta a questo interrogativo, il Metodo dell'Accettazione resta esposto all'accusa di essere un sistema arbitrario. Si tenga a mente come gli esempi da me riportati di Metodo dell'Accettazione supportino l'accettazione della Conclusione Ripugnante – producano ragioni in favore dell'aumento della popolazione. Ma questo esito non è affatto scontato: il Metodo dell'Accettazione potrebbe essere impiegato – ed è stato impiegato – per giustificare soluzioni diametralmente opposte, come il dovere di prodigarsi per l'estinzione del genere umano<sup>99</sup>. Chiamiamo questo problema “*Problema dell'Arbitrarietà*”.

D'altra parte il Metodo Intuizionista, così come formulato, presenta delle contraddizioni interne. Un primo problema è rappresentato dagli atteggiamenti di secondo livello – o, in generale, di livello superiore – che criticano alcuni dei nostri atteggiamenti di livello inferiore<sup>100</sup>. Poniamo il caso che un agente abbia, non due intuizioni pari-ordinate contraddittorie – del tipo “è giusto  $\phi$ ” e “non è giusto  $\phi$ ” –, ma una prima intuizione – “è giusto  $\phi$ ” – e una seconda intuizione di livello superiore che critichi la prima – “ignora l'intuizione per cui ‘è giusto  $\phi$ ’”. Secondo il Metodo Intuizionista sia il giudizio di livello inferiore sia il giudizio di livello superiore sono cor-

<sup>97</sup> L'obiezione del seguace del Metodo Intuizionista è quella che Arrhenius rivolge a Tännsjö in ARRHENIUS 2006b, 48 s.; la risposta che tira in ballo il pregiudizio specista è di Tännsjö medesimo (TÄNNSJÖ 2004, 229 s.)

<sup>98</sup> Questo non vuol dire che diverse persone si troveranno sempre d'accordo sugli esiti del Metodo Intuizionista. Infatti, può ben darsi che due soggetti trovino intuitivi giudizi diversi. Tuttavia, se questi due soggetti seguono il Metodo Intuizionista e arrivano a un disaccordo, sanno da cosa questo disaccordo è generato: è generato dall'aver risposte emotive completamente diverse.

<sup>99</sup> Pensiamo all'Estinzione (v.86). David Benatar, sostenitore dell'Estinzione, ha espressamente riconosciuto come la sua posizione sia in palese contrasto con il senso comune, aggiungendo come questa circostanza non scalfisca in alcun modo la solidità dei suoi argomenti. Infatti, il senso comune sarebbe fallace «although my conclusions may be counter-intuitive, the dominant intuitions in this matter seem thoroughly untrustworthy» e dovremmo accettare conclusioni contro-intuitive «when one has a powerful argument, based on highly plausible premises [...] but which is rejected merely because of primal psychological features that compromise our judgment, then the counter-intuitiveness of the conclusion should not count against it» (BENATAR 2006, 205, 207 ss.).

<sup>100</sup> FRANKFURT 1971.

retti – entrambi sono intuizioni dell’agente. Ma questo cosa vuol dire? Vuol dire che, alla fine, per il Metodo Intuizionista è come se i due giudizi fossero pari-ordinati? Oppure, dalla correttezza del giudizio di secondo livello deriva che, alla fine, dobbiamo neutralizzare il giudizio di primo livello, ossia neutralizzare una nostra intuizione, ossia accettare la sua negazione, accettare un giudizio contro-intuitivo? Per poter attrarre qualcuno, il Metodo Intuizionista deve essere rifinito rispondendo a questo interrogativo – *Problema delle Intuizioni di Secondo Livello*.

Questo problema se ne chiama dietro un altro. Poniamo di adottare una versione del Metodo Intuizionista per cui le intuizioni di primo livello, che sono bersaglio di critiche da parte di intuizioni di secondo livello, vengono neutralizzate – seconda alternativa del capoverso precedente. Per questa versione del Metodo Intuizionista si daranno conclusioni contro-intuitive che devono essere accettate. Ma questo è vero anche secondo il Metodo dell’Accettazione. Una volta introdotto il problema degli atteggiamenti di secondo livello, resta possibile distinguere tra i due metodi o l’uno si riduce all’altro? (*Problema della Riduzione*)

Un altro problema, attinente al Metodo Intuizionista, è rappresentato dal caso di un agente che trovi di per sé contro-intuitivo precisamente il fatto che la correttezza di un giudizio dipenda in modo così stretto dalla intuitività di quel giudizio – trovando intuitivo, piuttosto, che le valutazioni morali siano affette da pregiudizi. Questo agente, pur trovando contro-intuitivo il Metodo Intuizionista, potrebbe ritenerlo corretto, credervi. Tuttavia, il metodo in cui crede gli intima di seguire le proprie intuizioni e queste, nel suo caso particolare, condannano quello stesso metodo. Avremmo un metodo del ragionamento che, almeno in certi casi, prescrive il proprio abbandono (*Problema della Meta-intuitività*)<sup>101</sup>.

Infine, un problema generale. La contrapposizione tra Metodo Intuizionista e Metodo dell’Accettazione sembra avere la sua origine in due modi profondamente diversi di concepire il senso del ragionamento pratico. Per il Metodo Intuizionista, accade che noi siamo recettivi, sensibili, a certe circostanze e non ad altre – ad esempio molti di noi sono recettivi al benessere delle persone che esistono ma non sono allo stesso modo recettivi al benessere delle persone possibili; dato ciò che troviamo intuitivo e contro-intuitivo, dato ciò a cui siamo e non siamo recettivi, dato come siamo fatti, il ragionamento morale si risolve nel trovare soluzioni a incongruenze insite nella nostra contingente identità morale. Per il Metodo dell’Accettazione il fatto che noi siamo sensibili a certe cose e non ad altre non è sempre rilevante; al contrario, alcune delle nostre sensibilità sarebbero influenze distorsive che ci allontanano dalla soluzione corretta. Dovremmo liberarci di alcune delle nostre contingenze – in un modo da chiarire: ritorna il Problema dell’Arbitrarietà. Il ragionamento morale avrebbe, per il Metodo dell’Accettazione, una portata rivoluzionaria che manca nel Metodo Intuizionista. È un’alternativa davvero generale. Vogliamo che la morale vada dagli esseri umani o gli esseri umani dalla morale?<sup>102</sup> Chiamiamo questo problema “*Problema della contingenza*”<sup>103</sup>.

<sup>101</sup> Nella fase di referaggio è stato notato come il Metodo Intuizionista possa risultare in sé contro-intuitivo. Originariamente avevo ommesso ogni accenno a quello che ho chiamato “Problema della meta-intuitività”, ma la perplessità, comprensibile, del referee mi ha spinto a includerlo tra le questioni metodologiche poste dall’analisi del ragionamento morale. Altrove ho avanzato l’idea secondo la quale il metodo del ragionamento di coloro che criticano l’oggettività dei valori morali – ad esempio soggettivisti e nichilisti (*error-theorists*), qualora siano distinguibili – consista in una versione del Metodo Intuizionista a cui si aggiunge la regola per cui, quando la negazione dell’oggettività va contro le nostre intuizioni, sappiamo che le nostre intuizioni ci stanno illudendo. Il presupposto sarebbe che il critico dell’oggettività dei valori morali è qualificato dalle sue credenze – di stampo negazionista – sui valori morali, ma non anche dal fatto che si trovi a suo agio con questo apparato di credenze, ossia, non anche dal fatto che trovi perfettamente intuitive tutte le implicazioni delle sue credenze. Questo soggetto, nel momento in cui ha un’intuizione di stampo oggettivista, “si ricorda” che questa intuizione è soltanto un’illusione – e questo basta a fare di lui un critico dell’oggettività. In questo senso, il critico dell’oggettività accetta che si diano illusioni in ambito morale, ma limitatamente alle intuizioni in favore dell’oggettività dei valori.

<sup>102</sup> FRANKENA 1973, 116.

<sup>103</sup> Quando parlo di sussunzione di certi argomenti (quelli discussi tra §3 e §7) entro il Metodo Intuizionista bisogna tenere a mente due precisazioni. In primo luogo vi sono snodi argomentativi, sparsi nella letteratura, che effet-



Si tratta di problemi di metodo per nulla confinati nel dibattito di etica della popolazione<sup>104</sup> e che spesso vengono trascurati.

tivamente stridono con il Metodo Intuizionista. Ad esempio, Parfit nella sua difesa del Perfezionismo – §5 – puntualizza come sebbene le soluzioni da lui prospettate possano in sé sembrare strane, sono comunque quelle *meno implausibili* (PARFIT 2016, 120, 126). Per come da me introdotto il Metodo Intuizionista richiede che una soluzione morale, per meritare questo nome, non appaia strana affatto. Allo stesso tempo il criterio indicato da Parfit – la ricerca della soluzione meno implausibile – sembra assegnare rilevanza al senso di correttezza: il criterio non appare del tutto antitetico alle regole del Metodo Intuizionista. Avremo una sorta di regola ibrida posta a metà tra il Metodo Intuizionista e il Metodo dell'Accettazione. In secondo luogo, occorre precisare che la mia ricostruzione in termini psicologisti del concetto di intuizione potrebbe non essere accolta con favore dai vari autori che hanno proposto e discusso le varie soluzioni ai dilemmi dell'etica della popolazione (tra §3 e §7). Potrebbe ben darsi che molti di questi autori quando parlano di implicazioni contro-intuitive non vogliano tanto dire che certe implicazioni suscitano in noi una reazione emotiva di avversione, quanto che queste implicazioni sono irrazionali. Questi autori potrebbero allora tranquillamente concedere che alcune implicazioni che suscitano in noi reazioni emotive di avversione – conclusioni contro-intuitive in senso psicologico – debbano essere accettate. In tal caso, questi autori adotterebbero non il Metodo Intuizionista ma il Metodo dell'Accettazione e dovrebbero affrontare i problemi che ho menzionato. In particolare andrebbe loro chiesto cosa è che rende i giudizi da loro considerati intuitivi espressioni della razionalità umana e non soltanto espressioni di stati emotivi di approvazione o disapprovazione.

<sup>104</sup> Rinvio in proposito ad alcuni testi classici già citati in cui questi problemi di metodo si pongono (SINGER 1972, THOMSON 1985) e vengono analizzati (UNGER 1996, GREENE 2013).

*Riferimenti bibliografici*

- ARRHENIUS G. 2000a. *An Impossibility Theorem for Welfarist Axiologies*, in «Economics and Philosophy», 2, 2000, 247ss.
- ARRHENIUS G. 2000b. *Future Generations. A Challenge for Moral Theory*, dissertazione di dottorato, Uppsala.
- ARRHENIUS G. 2009. *Can the Person Affecting Restriction Solve the Problems in Population Ethics?*, in ROBERTS M., WASSERMAN D. (eds), *Harming Future Persons. Ethics, Genetics and the Non-Identity Problem*, Dordrecht – Heidelberg – London – New York, Springer, 289 ss.
- ARRHENIUS G. 2016. *Population Ethics and Different-Number-Based Imprecision*, in «Theoria», 82, 2, 2016, 166 ss.
- ARRHENIUS G., RABINOWICZ W. 2015. *The Value of Existence*, in HIROSE I., OLSON J. (eds.), *The Oxford Handbook of Value Theory*, Oxford, Oxford University Press, 424 ss.
- ARRHENIUS G., RYBERG J., TÄNNSJÖ T. 2017. *The Repugnant Conclusion*, disponibile in: <https://plato.stanford.edu/entries/repugnant-conclusion/> (consultato il 04/03/2020).
- BENATAR D. 2006. *Better Never to Have Been*, Oxford – New York, Oxford University Press.
- BLACKORBY C., BOSSERT W., DONALDSON D. 1997. *Critical-level Utilitarianism and Population Ethics Dilemma*, in «Economics and Philosophy», 13, 2, 1997, 197 ss.
- BOONIN D. 2014. *The Non-identity Problem & the Ethics of Future People*, Oxford, Oxford University Press.
- BRIGAGLIA M. 2018. *Genealogia della normatività. La normatività come controllo*, in «Diritto & Questioni Pubbliche», XVIII, 1, 2018, 59 ss.
- BROOME J. 2004. *Weighing Lives*, New York, Oxford University Press.
- CARLSON E. 1998. *Mere Addition and Two Trilemmas of Population Ethics*, in «Economics and Philosophy», 14, 2, 1998, 283 ss.
- CELANO B. 2013. *I diritti nello Stato costituzionale*, Bologna, il Mulino.
- ELSTER J. 1979. *Ulysses and the Sirens*, Cambridge, Cambridge University Press.
- FRANKENA W.K. 1973. *Ethics*, seconda edizione, Englewood Cliffs (NJ), Prentice Hall.
- FRANKFURT H.G. 1971. *Freedom of the Will and the Concept of a Person*, in «The Journal of Philosophy», 68, 1, 1971, 5 ss.
- GREAVES H. 2017. *Population Axiology*, in «Philosophy Compass», 12, 11, 2017, 1ss.
- GREENE J. D. 2013. *Moral Tribes: Emotion, Reason and the Gap between Us and Them*, New York, The Penguin Press.
- GRIFFIN J. 1986. *Well-Being. Its Meaning, Measurement, and Moral Importance*, Oxford, Oxford University Press.
- HUEMER D. 2005. *Ethical Intuitionism*, Basingstoke, Palgrave MacMillan.
- HUEMER D. 2008. *In Defence of Repugnance*, in «Mind», 117, No. 468, 2008, 900 ss.
- HURKA T.M. 1983. *Value and Population Size*, in «Ethics», 93, 3, 1983, 496 ss.
- MCMAHAN J. 1981. *Problems of Population Theory*, in «Ethics», 92, 1, 1981, 96 ss.
- MCMAHAN J. 2009. *Asymmetries in the Morality of Causing People to Exist*, in ROBERTS M., WASSERMAN D. (eds), *Harming Future Persons. Ethics, Genetics and the Non-Identity Problem*, Dordrecht – Heidelberg – London – New York, Springer, 49 ss.
- MCMAHAN J. 2013. *Causing People to Exist and Saving People's Lives*, in «The Journal of Ethics», 17, 1/2, 2013, 5 ss.
- MCTAGGART J.M.E. 1921. *The Nature of Existence*, Cambridge, Cambridge University Press.

- NARVESON J. 1967. *Utilitarianism and New Generations*, in «Mind», 76, 301, 1967, 62 ss.
- NARVESON J. 1973. *Moral Problems of Population*, in «The Monist», 57, 1, 1973, 62 ss.
- NG Y.K. 1989. *What Should We Do About Future Generations?*, in «Economics and Philosophy», 5, 2, 1989, 235 ss.
- PARFIT D. 1984. *Reasons and Persons*, Oxford – New York, Oxford University Press.
- PARFIT D. 1986. *Overpopulation and the Quality of Life*, in SINGER P. (ed.), *Applied Ethics*, Oxford, Oxford University Press, 145 ss.
- PARFIT D. 2011. *On What Matters. Volume I*, Oxford, Oxford University Press.
- PARFIT D. 2016. *Can We Avoid the Repugnant Conclusion?*, in «Theoria», 82, 2, 2016, 110 ss.
- PARFIT D. 2017. *Future People, the Non-Identity Problem, and Person-Affecting Principles*, in «Philosophy & Public Affairs», 45, 2, 2017, 118 ss.
- PELLEGRINO G. 2012. *Etica delle popolazioni: l'inevitabilità della qualità della vita e dei paradossi che ne conseguono*, in «Iride», XXV, 1, 2012, 27 ss.
- RAWLS J. 1971. *A Theory of Justice*, Cambridge (Mass.), Harvard University Press.
- SIDER T. R. 1991. *Might Theory X Be a Theory of Diminishing Marginal Value?*, in «Analysis», 51, 4, 1991, 265 ss.
- SIDGWICK H. 1907. *The Methods of Ethics*, settima edizione, Indianapolis-Cambridge, Hackett Publishing Company.
- SIKORA R.I. 1975. *Utilitarianism: The Classical Principle and the Average Principle*, in «Canadian Journal of Philosophy», 5, 3, 1975, 409 ss.
- SINGER P. 1972. *Famine, Affluence, and Morality*, in «Philosophy & Public Affairs», 1, 3, 1972, 229 ss.
- TÄNNSJÖ T. 1998. *Hedonistic Utilitarianism*, Edimburgo, Edinburgh-University Press.
- TÄNNSJÖ T. 2004. *Why We Ought to Accept the Repugnant Conclusion*, in RYBERG J., TÄNNSJÖ T. (eds.), *The Repugnant Conclusion*, Amsterdam, Library of Ethics and Applied Philosophy, vol. 15, Manging Editor, 219 ss.
- TEMKIN L.S. 1987. *Intransitivity and the Mere-Addition Paradox*, in «Philosophy & Public Affairs», 16, 2, 1987, 138 ss.
- TEMKIN L.S. 2012. *Rethinking the Good. Moral Ideals and the Nature of Practical Reasoning*, New York, Oxford University Press.
- THOMSON J.J. 1985. *The Trolley Problem*, in «The Yale Law Journal», 94, 6, 1985, 1395 ss.
- UNGER P. 1996. *Living High & Letting Die. Our Illusion of Innocence*, New York – Oxford, Oxford University Press.
- WILLIAMS B.A.O. 1973a. *The Makropulos Case: Reflections on the Tedium of Immortality*, in ID. *Problems of the Self*, Cambridge, Cambridge University Press, 82 ss.
- WILLIAMS B.A.O. 1973b. *A Critique of Utilitarianism*, in SMART J.J.C., WILLIAMS B.A.O. (eds.), *Utilitarianism for & against*, New York, Cambridge University Press, 77 ss.