

Rawls e il problema della stabilità

Roberto Brigati

1. Giustizia e bontà

Secondo la formula usata dal suo stesso ideatore, la teoria della giustizia di John Rawls è «una teoria come-se non-storica»¹. Questa dizione (resa ancor più goffa dalla traduzione) esprime il fatto che il processo di formazione della giustizia descritto da Rawls non avrebbe mai potuto verificarsi nella realtà. Pertanto la teoria di Rawls è di fatto un invito a un inusitato esperimento mentale. Non solo dobbiamo calarci in una situazione nomotetica originaria in cui una pluralità di attori sociali è impegnata a stabilire per contratto i principi fondamentali della giustizia distributiva; non solo dobbiamo assumere la totale parità iniziale di tali attori, assicurata dall'ipotesi per cui ciascuno ignora quali siano le proprie preferenze, capacità, interessi particolari e motivazioni. Dobbiamo anche stare al gioco di Rawls accettando la scommessa per cui i principi che *immaginiamo* essere scelti dai contraenti in quella «posizione originaria» *definiscono* la giustizia. Chiarisco subito che nel presente lavoro mi dispongo, per quanto possibile, a stare a questo gioco. Le obiezioni rivolte all'architettura della posizione originaria occupano ormai una letteratura di vaste proporzioni²; ma qui ipotizzerò che sia possibile raggiungere un «equilibrio riflessivo» (così Rawls) tra l'idea intuitiva della giustizia e quella che scaturisce dal dispositivo idealizzato di Rawls. Mi chiederò invece a quali condizioni metodologiche sia possibile portare fino in fondo l'esperimento in questione: cioè pensare come realizzata una società che sia (in tal senso) giusta, trasfor-

¹ J. RAWLS, *The Basic Structure as Subject*, in A.I. GOLDMAN - J. KIM (edd), *Values and Morals*, Dordrecht 1978, p. 70 n. 6. Si veda ora una nuova versione di questo saggio in *Political Liberalism*, New York 1993 (cfr. p. 273-74).

² A questo coro di critiche ho già avuto modo di aggiungere la mia voce in *Rawls e la teoria dell'equità*, in «Discipline Filosofiche», 1993, 1.

mando il contrattualismo trascendentale di Rawls in un contrattualismo reale.

È nella terza parte (dedicata ai «Fini») del suo opus *magnum*³ che Rawls mette mano alle condizioni di possibilità di una società ordinata secondo i principi della «giustizia come equità (*fairness*)». In numerosi articoli posteriori al 1971, Rawls si è dedicato, poi, sempre più al problema della stabilità politica del suo modello sociale (mitigando forse le ambizioni etiche della teoria)⁴. Porsi il problema della stabilità di una simile società ben-ordinata significa indagare le motivazioni dell'agire: passare cioè dalla moralità formale del contratto a quella concreta della convivenza seppure regolata. Per questo Rawls non può sfuggire a una discussione della *teoria del bene*. Ora, com'è noto, la giustizia come equità di Rawls rappresenta nella ricerca etica contemporanea una delle più forti prese di posizione in favore di un'impostazione *deontologica*. Ciò implicherebbe il far astrazione da qualunque considerazione finalistica – che farebbe dipendere la giustizia da un'immagine inevitabilmente parziale di ciò che «è bene per» l'umanità – e conseguentemente anteporre il concetto di «giusto» a quello di «bene» nell'ordine della spiegazione come in quello della fondazione. Ma la valenza di questa mossa corre forse il rischio di essere equivocata. Sembra di poter dire che per Rawls la precedenza è pratica, non logica: non vuol dire che la giustizia possa essere definita indipendentemente da qualunque idea del bene, ma piuttosto che, in caso di conflitto, dovranno in generale prevalere le esigenze della giustizia rispetto a quelle dell'utilità pubblica anche se acclarata. Senza questa interpretazione sarebbero inintelligibili molte delle istanze fatte valere da Rawls nell'elaborazione della sua dottrina. Fra i numerosi aspetti della teoria rawlsiana che dipendono essenzialmente da una «teleologia» più o meno sottintesa, terremo presenti soprattutto i tre seguenti: la compilazione da parte di Rawls di una lista di «beni primari» desiderabili, che rappresentano per così dire la causa materiale della giustizia distributiva; la dipendenza, ripetutamente

³ *A Theory of Justice*, Cambridge 1971; trad. it. a cura di S. MAFFETTONE, *Una teoria della giustizia*, Milano 1982, 1989³ (d'ora in poi TG; i numeri di pagina tra parentesi nel testo si riferiscono a questa traduzione).

⁴ Le novità introdotte nei testi più recenti non giustificano comunque una distinzione tra «due Rawls» (cui accenna S. MAFFETTONE, *Valori comuni*, Milano 1989, p. 251). Molti dei testi successivi a *Una teoria della giustizia* sono integrati e riveduti nel recente *Political Liberalism*, cit., che Rawls sembra ora considerare come la versione definitiva del proprio pensiero. Benché basato appunto su materiale già pubblicato, il volume fornisce nuove formulazioni e rettifiche, che richiederebbero forse un esame separato. In questo lavoro mi limiterò a sottolineare gli elementi di continuità rispetto alle dottrine precedenti.

affermata, tra l'architettura della posizione originaria e una certa «concezione della persona»; e la stessa decisione d'intraprendere una dettagliata teoria del bene, distinguendola nettamente in due tempi: un'idea minima del bene, in dotazione alle parti nella posizione originaria, e una teoria completa riservata a una società in cui siano già vigenti i principi della giustizia. Infatti nella posizione originaria si dovrà tener conto di una teoria parziale (*thin*) del bene, in grado di fornire le motivazioni minime per l'agire (tra esse, par di capire, la stessa preferenza per una società intesa come schema di cooperazione) e di chiarire il carattere *razionale* della scelta dei principi di giustizia. Una teoria completa darà invece per acquisiti i principi e procederà alla definizione delle azioni buone o valori, e del valore morale delle persone. In questa connessione acquista rilievo, in particolare, il «senso di giustizia» che i soggetti morali dovrebbero, secondo Rawls, sviluppare in una società ben-ordinata⁵: «le parti possono avere reciproca fiducia nel comprendere e nell'agire in conformità con qualsiasi principio sul quale si siano infine accordate» (TG, 132). Apparentemente questa fiducia non è determinata da alcun aspetto particolare della posizione originaria così come la definisce Rawls; è un elemento in più che assicura gratuitamente l'osservanza del patto, qualunque esso sia. Non che il senso di giustizia sia ciò che rende giusti i principi: questi sono giusti in sé, fondandosi solo sull'autonomia della ragione. Il senso della giustizia è necessario perché una società giusta si regga in piedi: è richiesto cioè per l'*enforcement* dei principi, e non affinché i principi stessi siano vincolanti. Questa necessità ha peraltro un carattere non solo empirico, ma schiettamente concettuale: in altri termini, la presenza di un senso di giustizia diffuso e riconosciuto rientra nella *definizione* di una società giusta⁶. Senza questo elemento, tutto il dispositivo contrattuale resterebbe nel cielo delle astrazioni teoriche.

Le *quaestiones iuris* che Rawls si trova ad affrontare sono quindi due.

1. Come giustificiamo i principi della giustizia?
2. Come giustificiamo la necessità di rispettarli? La risposta di Rawls al primo problema è improntata a una sorta di deduzione

⁵ Cfr. *Una teoria della giustizia*, cap. VIII. Per il concetto di società ben-ordinata si veda in particolare il § 69; in *Political Liberalism* Rawls sembra considerare «irrealistica» la nozione di società ben-ordinata impiegata in precedenza.

⁶ Su questo aspetto è fondamentale il lungo articolo *Kantian Constructivism in Moral Theory*, in «Journal of Philosophy», 1980, pp. 515-72 (ora ricompreso nei primi tre capitoli di *Political Liberalism*, cit.).

trascendentale: in un certo senso, l'esperimento mentale di cui si diceva è infatti un complesso tentativo di dimostrare (si direbbe quasi ostensivamente) la giustezza (*rightness*) della giustizia. Questa deduzione ci assicura che il senso di giustizia protegge principi giusti e morali piuttosto che iniqui o immorali. Ma anche il secondo problema è risolto trascendentalmente sulla scorta dell'impostazione kantiana. *Pacta sunt servanda*: l'inosservanza universale di questa clausola distruggerebbe l'attività stessa del patteggiare. Il senso di giustizia, come vedremo, rientra non solo nella definizione di società giusta, ma anzitutto nella definizione della «persona» funzionale a tale società. Perché dunque Rawls insiste tanto sulla specificazione genetica del senso di giustizia, al punto di dedicarvi un intero capitolo di *Una teoria della giustizia*? La ragione di ciò è che Rawls ha bisogno di trasformare la *quaestio iuris* in una *quaestio facti*. Non si tratta solo di comprendere che è (trascendentalmente) giusto obbedire ai principi, ma di garantire la sussistenza di una società giusta. O, detto con termini che saranno chiari tra poco, non si tratta solo di capire perché è «kantianamente» razionale rispettare i patti, ma di fondare la razionalità «hobbesiana» di tale osservanza.

Questa concezione affida al senso di giustizia due compiti distinti, di cui Rawls non coglie la tensione. Per un verso il senso di giustizia dovrebbe rappresentare una motivazione per l'agire morale, ponendo le basi per una strategia concreta di risposta sociale all'egoismo; per un altro, esso contribuirebbe alla risposta teorica alla domanda su che cosa sia la giustizia. Una società è giusta solo se in essa vige un senso di giustizia. Ma un argomento trascendentale non può fornire la motivazione dell'azione: è questo, evidentemente, il nucleo della critica hegeliana al formalismo morale kantiano. Insomma per Rawls il senso di giustizia rientrerebbe sia nell'ordine delle cause che in quello delle giustificazioni dell'agire morale. Per questo la sua teoria partecipa di taluni caratteri delle teorie deontologiche come di quelle teleologiche.

2. Geometria del bene

La definizione «parziale» di bene fa leva sul concetto di razionalità: «A è un buon X se e solo se A possiede le proprietà che è razionale volere in un X» (TG, 330). In questa fase originaria non vige ancora una razionalità piena; stando ai presupposti di Rawls, la padronanza del sillogismo pratico da parte del singolo dovrebbe essere sufficiente alle esigenze della teoria parziale. Pertanto la razionalità in questo senso coincide con la *Zweck-*

mäßigkeit, la ragione puramente strumentale: «Sullo sfondo è sempre presente il punto di vista dal quale si analizza un manufatto, una parte funzionale, un ruolo» (TG, 333). Infatti «il punto di vista dal quale le cose vengono giudicate buone o cattive non ha niente di necessariamente giusto o moralmente corretto» (*ibidem*). Naturalmente questa mossa, se allontana temporaneamente Rawls dal tracciato «kantiano», ripetutamente invocato, consente al suo contrattualismo di godere del vantaggio metodologico principale della tendenza «hobbesiana» in sede di fondazione dell'etica. In tal modo difatti non occorre postulare che le inclinazioni fondamentali degli uomini siano buone. È sufficiente che, date tali inclinazioni (o meglio, data la consapevolezza di avere delle inclinazioni, anche se non si sa quali siano: si veda più sotto la nozione del «velo d'ignoranza»), le persone si pongano in modo razionale nei confronti di esse: cioè vogliano essere liberi di perseguirle nei limiti del possibile. In conseguenza di ciò, le parti nella posizione originaria sceglieranno principi tali da promuovere questa libertà, ma tali anche da prevenire rimostranze o giustificate opposizioni da parte degli altri (la cui conseguenza potrebbe essere il *bellum omnium*): «un accordo su questi principi è il miglior modo che ciascuno ha per garantire i suoi fini in relazione alle alternative a disposizione» (TG, 111).

Vediamo di ricostruire le linee di questa «geometria morale» (TG, 113). La teoria parziale del bene si occupa soltanto di ciò che è *zweckrational*: il *wertrational* è lasciato alle preferenze individuali (messe fra parentesi nella posizione originaria). Ne segue che il buono (in senso *thin*, e quindi *zweckrational*) non è immediatamente il giusto: «Il concetto di razionalità non è di per sé stesso una base adeguata per il concetto di giusto; e nella teoria contrattualista quest'ultimo è derivato in altro modo» (TG, 334). Questo «altro modo» è appunto l'esperimento mentale della «posizione originaria»; ma per gli scopi di questo lavoro, risulta in primo luogo significativo il fatto che la razionalità del bene (o bontà come razionalità), intesa in chiave disposizionale, non conduce a esiti diversi da quelli a cui si giunge attraverso la fondazione contrattualista: «combinando insieme il disinteresse reciproco e il velo d'ignoranza, si ottiene lo stesso effetto determinato dalla benevolenza. Questa combinazione di condizioni, infatti, costringe ogni individuo nella posizione originaria a tenere conto del bene altrui»⁷. La tendenza ad agire moralmente

⁷ *Una teoria della giustizia*, 134. Il «velo d'ignoranza» è la metafora che esprime lo stratagemma di privare i contraenti di tutte le informazioni riguardanti i propri interessi peculiari e fini ultimi.

collima con la scelta più razionale che si potrebbe fare nella posizione originaria in vista della protezione dei propri interessi. A questo punto non ci sono più obiezioni all'esercizio di una razionalità puramente strumentale. Nella posizione originaria non risulta necessario tener conto di valori separati dai piani di vita degli individui: la teoria della giustizia è soltanto l'ostensione della strategia più razionale per realizzare questi piani. In altre parole la giustizia è il modo più razionale di perseguire il bene; una conclusione che riconcilia trionfalmente il piano deontologico e quello teleologico, e riunisce i meriti della linea hobbesiana e di quella kantiana.

C'è quindi una duplicità nell'argomento di Rawls⁸. A prima vista sembra inserirsi – con un apporto originale rappresentato dalla struttura contrattualista – nella linea classica dell'egoismo razionale hobbesiano. Sembra insomma che per Rawls l'altruismo possa scaturire da un egoismo avveduto; ossia che dalla razionalità strumentale possa nascere spontaneamente la razionalità piena. Invece si scopre che occorre un elemento addizionale, un dato «kantiano», appunto il senso di giustizia; un po' come l'utilitarismo alla Bentham non può far a meno di un principio di benevolenza come collante sociale⁹. Rawls deve allora dimostrare che una società ben-ordinata è tale da far sorgere negli individui il senso della giustizia nel modo più spontaneo e sicuro (naturalmente debbono esserci altri modi, visto che molti individui hanno un vivo senso della giustizia pur vivendo in società largamente ingiuste). A questo punto la geometria morale di Rawls abbandona la linea trascendentale: occorre una dimostrazione empirica, ed è qui che la teoria evidenzia una lacuna.

Prima di esaminare questo nucleo empirico, occorre aggiungere alcuni elementi ulteriori al quadro della teoria parziale del bene. Ci potrebbe essere infatti un'obiezione radicale che bloccherebbe ogni sviluppo empirico del razionalismo morale (e non solo rawlsiano). L'obiezione proviene dal bagaglio culturale del pessimismo classico, e si concentra sull'incapacità dell'uomo di conoscere il bene, anche se inteso in senso parziale come «il proprio bene». I «beni primari», che ognuno dovrebbe desiderare, sarebbero dunque un elenco caotico di preferenze totalmente

⁸ La duplicità è correttamente evidenziata da B. BARBER, *Justifying Justice: Problems of Psychology, Politics and Measurement in Rawls*, in N. DANIELS (ed), *Reading Rawls*, Oxford 1975, pp. 315 ss.

⁹ Un'analogia notata da B. BARRY, *The Liberal Theory of Justice. A Critical Examination of the Principal Doctrines in «A Theory of Justice by John Rawls»*, Oxford 1973, p. 24.

idiosincratice, in cui non è possibile rinvenire una qualsivoglia regolarità. Si potrebbe allora mettere in dubbio la possibilità stessa di vivere secondo ragione; se la natura umana (o la natura in generale) è intrinsecamente irrazionale, la pretesa di individuare anche solo un ordinamento *thin* delle preferenze individuali è vana. Il bene di una persona sarebbe così indeterminabile. Ma nel quadro di una teoria della scelta razionale, il concetto di (bontà come) razionalità non è un *primum* inanalizzabile. Intanto può essere ulteriormente delucidato definendo alcuni principi deliberativi elementari (efficacia, inclusività, ecc.). Non è chiaro se Rawls pensi di poter risolvere analiticamente il concetto, senza residuo, in un insieme di concetti più primitivi. Certamente questo *explaining away* gioverebbe alla dimostrazione *more geometrico* vagheggiata da Rawls. In ogni caso il corso d'azione idealmente razionale è quello che la persona interessata sceglierebbe se avesse a disposizione tutte le informazioni rilevanti. «In breve, il nostro bene è determinato dal piano di vita che adotteremo con completa razionalità deliberativa se il futuro fosse accuratamente previsto e adeguatamente realizzato nell'immaginazione» (TG, 347). Nella vita concreta, è possibile quindi non sapere perfettamente quale sia davvero il nostro bene. Ma questo non implica il fallimento della razionalità deliberativa. Essa si esplica in negativo, in un senso peculiarmente minimalista: «Agire con razionalità deliberativa può solamente garantirci che la nostra condotta non merita biasimo» (TG, 348)¹⁰. La razionalità non possiede alcuna garanzia intrinseca di bontà se non in negativo; potremmo dire: anche se il mondo è intrinsecamente irrazionale, nessuno potrà biasimare il mio comportamento per non essere stato irrazionale.

L'altro importante elemento (non del tutto disgiunto da quello che abbiamo appena visto) che va aggiunto al quadro è di natura psicologica: si tratta di quello che Rawls chiama il «principio aristotelico», secondo il quale «gli esseri umani provano piacere nell'esercitare le loro capacità effettive ... e il loro piacere aumenta via via che la capacità si realizza o cresce la sua complessità» (TG, 351)¹¹. Seppur con cautela, il principio è presentato da

¹⁰ Tale garanzia deve valere per ipotesi: «... un individuo razionale deve sempre agire in modo da non doversi mai biasimare indipendentemente da come vanno a finire le cose» (TG, 348).

¹¹ Rawls non insiste sulla paternità aristotelica del principio. La formulazione più prossima che ho potuto trovare è in *Etica Nicomachea*, 1174b, 18-20: «in ogni cosa dunque l'attività migliore sarà quella che si esplica quando si ha la miglior disposizione verso ciò che è l'oggetto migliore che cade sotto di essa»

Rawls come un fatto empirico, probabilmente risalente a un dato biologico-evolutivo (TG, 355). Benché possa apparire come un elemento di scarsa incidenza, non va sottovalutata l'importanza di questo «principio»: è infatti l'unico punto in cui Rawls viene alle prese direttamente col problema della felicità. La felicità è possibile: esiste – per usare un vocabolario della tradizione pessimistica – uno spazio tra il dolore e la noia, ed è lo spazio dell'esercizio e dello sviluppo attivo delle nostre facoltà. Affrontando il problema del contenuto del bene, dunque, l'etica contrattualista deve recuperare un elemento di etica «perfezionista», precedentemente rigettata¹². Ma il recupero rimane a mezza strada. Se il principio fosse una legge universale, se cioè rientrasse nella definizione di razionalità¹³, tutto il castello dell'etica rawlsiana risulterebbe immediatamente superfluo. Dunque, il principio «esprime una tendenza e non uno schema invariabile di scelta» (TG, 353). Se è così, però, esso non può veramente servire a un ordinamento delle preferenze individuali secondo il valore. La definizione del bene resta «puramente formale» (TG, 349): un involucro il cui riempimento può venire solo dalle inclinazioni individuali e non può essere descritto in anticipo, come si vede dal bell'esempio «spinoziano» dato da Rawls: «Immaginate, quindi, un individuo il cui unico piacere sia quello di contare i fili d'erba ... [Questo piacere] sarà per lui il fine che regola il programma delle sue azioni, e questo dimostra che è buono per lui» (TG, 356)¹⁴. Rawls si trova allora stretto fra due opzioni: o ammettere che i fini ultimi sono irriducibilmente e irrazionalmente individuali, rinunciando a una teoria della bontà come razionalità, oppure accettare come razionale qualunque inclinazione personale, col rischio di non poter più giustificare alcuna prescrizione etica.

3. La genesi della moralità

«Una concezione di giustizia è più stabile di un'altra se il senso di giustizia che tende a ispirare è più forte e più adeguato a

(trad. Plebe). Cfr. anche 1174a, 1 ss.: «Infatti nessuno sceglierebbe di vivere mantenendo per tutta la vita la mentalità di un fanciullo, godendo di ciò che godono soprattutto i fanciulli».

¹² Cfr. *Una teoria della giustizia*, § 50. Perfezionista è la concezione che identifica il bene con l'eccellenza.

¹³ L'ipotesi è di B. BARRY, *The Liberal Theory*, cit., p. 28.

¹⁴ C'è un esempio, più drammatico, in Spinoza: «se qualcuno si accorgesse che potrebbe vivere più a suo agio infisso in croce che a tavola, sarebbe il più stolto degli uomini se non si facesse crocifiggere» (*Epistolario*, XXIII, trad. it. a cura di A. DROETTO, Torino 1951, p. 151).

superare le inclinazioni disgregatrici» (TG, 373). Ma questo senso negli uomini non è innato; perciò occorrerebbe, a prima vista, un'indagine della sua genesi psicologica a partire da leggi generali. Per comprendere il modo in cui Rawls si pone il problema della stabilità, bisogna tener presente l'uso trascendentale che egli fa del concetto di società ben-ordinata, e parallelamente del concetto di persona. Rawls non sta semplicemente dicendo che le persone possiedono un senso di giustizia, di origine misteriosa, e che chi voglia edificare una società ben-ordinata è autorizzato a far assegnamento su questa caratteristica generale dell'uomo. Se così fosse il suo argomento sarebbe ingenuamente ottimistico, e difficilmente un'indagine empirica seria potrebbe supportarlo. In realtà egli parte da un ideale sociale, e costruisce una definizione della giustizia in grado di condurre il più vicino possibile ad esso: «La giustizia come equità è strutturata in modo da accordarsi con questa idea di società» (TG, 372). Più esattamente, gli ideali fondanti sono due: accanto a quello sociale ve n'è uno antropologico, una concezione della persona. Infatti i membri di una società ben-ordinata sono «persone morali libere ed eguali»¹⁵. Si tratta di una concezione «piena» della persona, che si distacca con chiarezza da quella «sparsa» dell'intuizionismo morale e del razionalismo classico. Mentre quest'ultimo richiede poco più che «l'idea del sé come soggetto che conosce»¹⁶, Rawls ambisce a caratterizzare la persona come immersa *per definizione* in una società e in un reticolo di concezioni, rapporti e principi morali. Ma attenzione ora all'impiego fatto da Rawls di questi due concetti, quello di persona e quello di società. Non si tratta di concetti empirico-fattuali; ancora una volta, Rawls non sta dicendo che l'uomo è così, o che ci sono buone speranze che possa diventarlo: altrimenti la sua tesi andrebbe suffragata da un poderoso apparato di indagine antropologica sul campo. Piuttosto, Rawls elabora un tipo umano ideale, che coincide in larga parte con quello del cittadino *spoudatos*, del libero individuo animato dalle virtù civiche classiche¹⁷. Questo ideale non è la premessa, ma l'idea guida (ossia l'idea in senso kantiano) della costruzione etico-politica. Non importa sapere se questo tipo sia

¹⁵ Cfr. per esempio *Kantian Constructivism in Moral Theory*, cit., p. 521.

¹⁶ *Political Liberalism*, p. 92.

¹⁷ La presenza di un ideal-tipo nella concezione di Rawls è notata da H.L.A. HART (*Rawls on Liberty and Its Priority*, 1973, ora in N. DANIELS [ed], *Reading Rawls*, cit., p. 252): «L'ideale è quello di un cittadino di ispirazione civile, che stima l'attività politica e il servizio verso gli altri come beni fra i maggiori della vita».

incarnato realmente nel mondo; importa che, per giungere alla giustizia (naturalmente alla giustizia nel senso di Rawls, cioè come *fairness*), è necessario presupporre un tipo antropologico come quello e, collegato ad esso, un ideale sociale in grado di ospitarlo e farlo prosperare. In breve, Rawls non scopre il suo ideale etico-politico, ma lo costruisce (e «costruttivismo» è l'etichetta che egli è giunto a scegliere per la sua teoria). Ma allora, ci si può chiedere, dove sta la cogenza argomentativa della sua teoria? Come si distingue da un puro *wishful thinking*? Ebbene, il trucco consiste nel fatto che Rawls in effetti non ha inventato le sue nozioni teoriche di sana pianta, anche se la teoria procede «come se» così fosse. In realtà egli ha ricavato la sua concezione esplicitando e mettendo a nudo quella che, a suo dire, è la coscienza collettiva delle democrazie occidentali in cui viviamo. Anche se non si presenta come un'indagine sull'uomo e sul mondo in cui esso vive, la teoria di Rawls è dunque inseparabile dal mondo in cui viviamo. Sul piano teorico Rawls va in cerca di un tipo ideale; ma su un piano che si potrebbe chiamare metateorico, egli vi fa *appello*, sicuro di trovare una rispondenza nei suoi lettori (e su questa possibilità si gioca la sua nozione di «equilibrio riflessivo»). Il discorso filosofico di Rawls, insomma, non solo *parla di* un tipo umano, ma si *rivolge a* questo tipo umano o a coloro che sono interessati ad approssimarvisi¹⁸.

Ciò nonostante, a fianco di questo procedimento costruttivo, sopravvive in Rawls (e anche nei testi posteriori al 1971) una traccia di ricerca empirica minima, un ritratto dell'uomo così come esso è o potrebbe essere. Se è vero, infatti, che Rawls si appella alla nostra coscienza civile, non può sottrarsi del tutto all'impegno di mostrare come questa coscienza ha potuto stabilirsi. Circa la genesi dei sentimenti morali, Rawls mette a confronto due tradizioni principali. La linea empirista (da Hume a Sidgwick) sostiene che non ci sono reali motivi per adottare una condotta giusta (cioè socialmente utile); le motivazioni vengono quindi fornite con l'educazione, e questo già nei primissimi anni di vita¹⁹. La teoria razionalista (ovviamente legata al nome di

¹⁸ La circolarità di questo appello è un elemento ulteriore di affinità con Aristotele, se è vero, come afferma Mario Vegetti, che l'etica aristotelica si rivolge «a chi già possiede la virtù, per averla appresa non teoricamente ma attraverso i dispositivi dell'educazione sociale quotidiana ... Il *gentleman* ateniese del IV secolo è certamente il destinatario ultimo, oltre che il protagonista, del discorso etico di Aristotele» (*L'etica degli antichi*, Roma - Bari 1989, p. 165).

¹⁹ Rawls sottolinea come la teoria freudiana del Super-Io s'inserisca sostanzialmente in questa tradizione (TG, 376-77).

Rousseau, ma anche, secondo Rawls, di Kant) sostiene invece che il senso morale deriva dallo sviluppo naturale delle «nostre innate capacità intellettuali ed emotive» (TG, 377) e si acquisisce pertanto con la maturazione e la comprensione del proprio ruolo nella società. Dal canto suo Rawls descrive tre stadi dell'apprendimento morale. La moralità autoritaria è quella che viene inculcata nel bambino dall'esempio e dalla guida dei genitori. Qui si accoglie un aspetto della tradizione empirista: «Il succedersi delle generazioni e la necessità di insegnare gli atteggiamenti morali ... ai bambini è una delle condizioni della vita umana» (TG, 379). La moralità associativa si raggiunge col passaggio al piano più ampio delle associazioni coi propri simili. Le virtù specifiche e più apprezzate in questo contesto sono quelle relative alla cooperazione. L'individuo comincia a imparare a identificarsi con la prospettiva altrui e a provare sentimenti di colpa nei confronti degli altri membri dell'associazione. Infine la moralità dei principi subentra con la piena maturazione delle capacità intellettuali e la comprensione degli standard morali validi in generale (cioè, sembra di poter dire, validi in ogni associazione possibile o forse nella più ampia delle associazioni possibili). Questi standard non sono che i principi di giustizia, i quali vengono ora desiderati di per sé e non più relativamente a una particolare cerchia di amici o di partner. Con questo stadio Rawls si riallaccia evidentemente alla tradizione razionalista: «I nostri sentimenti morali dimostrano di essere indipendenti dalle circostanze accidentali del nostro mondo, e il significato di tale indipendenza è dato dalla descrizione della posizione originaria e dall'interpretazione kantiana» (TG, 389). Nondimeno anche lo stadio dei principi, come i primi due, è formulabile a detta di Rawls come una «legge» psicologica (TG, 401).

Se è vero che Rawls ha l'ambizione di combinare la linea empirista e quella razionalista (TG, 378), bisogna anzitutto notare che i problemi delle due impostazioni sono speculari. Gran parte delle obiezioni cui va incontro la rudimentale psicologia morale proposta in *Una teoria della giustizia* derivano da qui. L'approccio empirista ha ovviamente un problema di legittimazione razionale: con che *diritto* la società trasmette ai giovani certi valori e non altri? D'altro canto il «buco» della teoria razionalista è di natura empirico-fattuale: come dimostrare che il libero sviluppo delle inclinazioni naturali conduce davvero alla moralità? Per Rawls, a quanto pare, i due problemi si elidono l'un l'altro, in quanto «il sentimento di giustizia non è un desiderio diverso da quello di agire in base ai principi che individui razionali sceglierebbero in una situazione iniziale che offre a ciascuno

eguale rappresentanza come persona morale» (TG, 391). Ma non mi pare che questa soluzione sia più che una risposta postulativa, una scommessa suffragata unicamente, per un verso, da un complicato esperimento mentale, per l'altro, da una serie di «leggi» psicologiche a dir poco dubbie.

Va detto che, dopo *Una teoria della giustizia*, Rawls non ha quasi più insistito sulla psicologia morale delineata nel testo che l'ha reso celebre. Tuttavia questo silenzio da parte sua non fa che rimarcare ancor più l'esigenza posta poc'anzi: infatti, anche nella riflessione sul costruttivismo politico, non sono affatto venuti a cadere i presupposti che rendevano necessario il ricorso a strumenti psicologici. In *Political Liberalism*, Rawls ribadisce la sua convinzione che l'umanità può raggiungere un ordine morale facendo conto unicamente sulla «natura umana» stessa. L'azione formativa dello stato è considerata (al pari dell'esperienza religiosa della rivelazione) un'istanza «esterna» a tale natura, mentre ovviamente la ragione è collocata all'«interno» di essa. Nello stesso tempo la concezione della psicologia morale è espressamente definita «filosofica, e non psicologica»²⁰, nel senso che Rawls rivendica un'autonomia della filosofia nel proporre ideali che gli individui *possono* trovare attraenti. Gli ideali devono quindi rispondere unicamente a un requisito di compatibilità con la struttura psicologica umana così com'è rilevata dall'indagine antropologica empirica. Un ideale politico integrato – sembra dire Rawls – in parte sfrutta le risorse già esistenti nella natura umana, e per il resto *plasma* la personalità psicologica che gli occorre. Questo può risolvere il problema della stabilità (se non quello della legittimazione, che come s'è detto e ripetuto pertiene a un altro ambito)? La verità, mi sembra, è che l'autonomia propositiva della filosofia politica non risolve il problema della stabilità, ma offre semmai una cornice in cui per la prima volta esso può essere posto come problema. La chiave è nella funzione d'appello di cui s'è detto. La stabilità del sistema dipende dalla sua capacità di chiamata: una chiamata alla responsabilità che però è diretta alla collettività in quanto insieme di individui modellati su quel sistema di ideali. In questa luce, si vede allora che la distinzione netta tra moralità associativa e moralità dei principi non regge: il rispetto dei principi non è separabile dal fatto che sono i principi della (più ampia) associazione di cui si è partecipi. Se dunque c'è un fatto che lascia sperare a Rawls di vincere la sua scommessa, esso non si trova sul piano delle

²⁰ *Political Liberalism*, cit., p. 86.

argomentazioni teoriche, ma nella chiusura metateorica di queste: appellandosi al cittadino consapevole, razionale e autonomo di una democrazia liberale in senso moderno, Rawls si assicura ad un tempo, *de iure*, il diritto a proporre una ricostruzione razionale della giustizia e, *de facto*, il consenso generale a questa ricostruzione. È vero dunque, in un certo senso, che solo *vivendo* in una società ben-ordinata si può essere giusti (nel senso di Rawls); ma questo non in forza di leggi psicologiche²¹. Il fatto è che la teoria di Rawls è una ricostruzione dei principi di una democrazia liberale, e solo chi appartiene a una tale società e sente di appartenervi è in grado propriamente di riconoscersi in una tale ricostruzione: *sa che cosa vuol dire* giustizia come equità.

4. Il problema del 'free-rider'

In larga misura, la teoria morale di Rawls si risolve in un tentativo di mediare lo scarto tra egoismo e altruismo. L'architettura metodologica sopporta buona parte del peso della mediazione. Quando la situazione iniziale è definita col metodo del velo d'ignoranza, l'egoismo (delle motivazioni) dà come risultato l'altruismo (dei principi). Perché questo altruismo si traduca anche sul piano delle opere, occorre invece una mediazione psicologica, un condizionamento. La seconda mediazione si rende necessaria, per Rawls, in quanto «in generale, agire con equità non è la risposta migliore di ciascuno al giusto comportamento dei suoi associati» (TG, 406). È sempre possibile che l'individuo si comporti come un *free-rider*, considerandosi al di sopra della legge e sfruttando l'equità degli altri senza sottoporsi ai doveri di cittadinanza²². Costui dunque effettua un ritorno all'egoismo sul piano strategi-

²¹ Naturalmente ciò non esclude l'azione del condizionamento psicologico; e l'idea sembra conforme alla concezione aristotelica della virtù morale come risultato dell'abitudine: *Eth. Nic.*, 1103b 4.

²² Una definizione del concetto di *free-rider* non rientra nei nostri scopi. Di recente R. TUOMELA - K. MILLER (*We-Intentions, Free-Riding, and Being in Reserve*, in «Erkenntnis», 36, 1992, pp. 25-52) hanno proposto una classificazione a seconda del danno inflitto dal *free-rider* alla società. Gli autori però ritengono essenziale che il *free-rider* consideri proprio dovere cooperare (p. 40); altrimenti, par di capire, sarebbe semplicemente un cittadino male informato. Mi pare, come minimo, che questo criterio non riesca a cogliere la vasta gamma delle sfumature dell'ipocrisia. Il caso più interessante dovrebbe essere proprio quello «superomistico»: un individuo che ritiene di non essere tenuto a contribuire al bene comune, pur godendone. È qui che si gioca la razionalità della morale e la possibilità di fornirne un criterio indipendente dalla *Zweckrationalität*.

co, dopo aver accettato la giustizia in funzione meramente tattica.

Il problema del *free-rider* è particolarmente increscioso in quanto inficia non tanto le teorie morali tradizionaliste basate sulla *pietas* o sul timor di Dio, ma proprio le teorie fondazioniste «hobbesiane», basate su una razionalizzazione dell'egoismo. La risposta assolutista di Hobbes contiene un tasso di coercizione che Rawls non è disposto a tollerare; ma c'è da chiedersi se la sua risposta psicologista sia più convincente. Lo status di legge psicologica attribuito ai fattori sociali che condizionano l'insorgere del senso di giustizia nell'individuo, come ho detto, appare dubbio. Ma c'è di più: la *Zweckrationalität* del *free-rider* può almeno pretendere lo stesso livello di necessità causale (nomologica) delle leggi adottate da Rawls. Rawls dovrebbe allora dimostrare che il «condizionamento alla giustizia» è o può essere (per lo più) più forte della tentazione «superomistica». Ora, una simile dimostrazione si preannuncia piuttosto ardua: bisognerebbe considerare un numero significativo di casi particolari di egoismo strategico, provando in ciascun caso che, per esempio, il comportamento in questione è riconducibile a (o almeno è costantemente accompagnato da) una storia personale di frustrazione, o scarso affetto, cattivi esempi, ecc. E anche se questa dimostrazione anamnesticamente fosse possibile, c'è da chiedersi se sia auspicabile. Di fronte al problema in questione, la concezione riassunta nelle tre leggi rawlsiane fa risaltare un'idea di apprendimento morale che assomiglia troppo alla «morale da schiavi» per risultare accattivante. Per inciso, Rawls avverte chiaramente quest'ultimo problema. Egli si chiede, infatti, se l'origine psicologica dei sentimenti morali non ne infici la validità («Immaginiamo allora che qualcuno faccia esperienza dei suggerimenti del proprio senso morale sotto forma di inibizioni inesplicabili che per il momento non è in grado di giustificare. Per quale motivo non dovrebbe considerarle soltanto come pulsioni nevrotiche?», TG, 420). La sua risposta non può che andare a parare sull'esplicabilità di quelle «inibizioni»: ancora una volta, si tratta di dare una legittimazione razionale. I molteplici vantaggi di un comportamento giusto in una società giusta possono essere non solo inculcati, ma *spiegati*, cioè illustrati contando sul potere della parola. Rispetto a questa risposta possono essere sollevati due ordini di problemi. Il primo è estremamente delicato e coinvolge un difficile argomento di psicologia cognitiva: c'è davvero una differenza essenziale tra le due modalità di apprendimento, quella del condizionamento e quella della persuasione? Se convincere vuol dire «riordinare gestalticamente» il campo cognitivo di un'altra persona, questo spostamento di informazioni è poi così diverso da una «manomis-

sione»? Riflettiamo ad esempio su questo: in linea di principio, lo stesso effetto di riordinamento gestaltico potrebbe anche essere ottenuto chirurgicamente o chimicamente? Non intendo però discutere questo punto. L'altro punto è più diretto: l'argomentazione razionale in favore di una società giusta può davvero convincere un *free-rider*? Se così fosse, pare che il problema del *free-rider* non sussisterebbe, o sarebbe solo un problema transitorio, dinanzi al quale non si richiederebbe altro che *calamus sumere et calculare*. Entro un certo limite, penso che Rawls tenda a concepire il problema dell'egoismo appunto in questi ultimi termini; ma non vedo ragione di condividere questo ottimismo.

Il fatto è che il formalismo morale di Rawls non gli permette di uscire dalla strettoia tra il postulato psicologico del senso di giustizia e la soluzione coercitiva di Hobbes²³. Una soluzione «geometrica» al problema del *free-rider* non è in vista. Ma proprio per questo il problema può essere affrontato solo sul piano della moralità associativa (a cui Rawls non vuole evidentemente ricorrere per i risvolti «comunitari» ed extra-razionali che comporta). Si direbbe anzi che la moralità associativa esista soltanto perché esiste il problema del *free-rider*. L'egoista può essere messo di fronte alla sua adesione esplicita o implicita a certi principi; gli si può rinfacciare la sua richiesta d'ammissione al club dei giusti. Non è possibile convincere il *free-rider*, ma in un certo senso non è neppure necessario: egli infatti rappresenta una figura logicamente parassitaria, che esiste soltanto perché esiste una comunità che vive in modo giusto. Il sabotatore auto-interessato deve aver caro il buon ordine della società, che gli garantisce la possibilità di continuare a perseguire il proprio tornaconto. Così l'interferenza creata dal *free-rider* non si «espande», rimane locale. Ed è proprio il fatto che il fenomeno del *free-rider* non può essere eliminato analiticamente a dare alla comunità il diritto di sanzionarlo praticamente. Il *free-rider* è un *problema* per la società, e lo rimane costantemente, rendendo necessaria un'attenta sorveglianza sociale. Se questa linea, brevemente accennata, è corretta, allora l'egoismo va accettato come un fatto naturale, ma proprio ciò fornisce una legittimazione alla *pratica* della giustizia: una pratica in qualche modo surrogatoria, legittimata per provvedere, *faute de mieux*, alle necessità poste dalla natura. L'egoismo

²³ Naturalmente esistono altre posizioni, come quella, che Rawls collega al nome di Ross (TG, 391), per cui il senso morale è un «puro atto di coscienza» del tutto gratuito. Ma uscire da uno schema stimolo-risposta, per spiegare l'apprendimento morale, non porta a nessun risultato concreto.

rimane una scelta permanentemente aperta, di fronte alla quale l'ideale fondante della società giusta può mostrare, se ne è capace, la propria bontà. Il problema, in altri termini, si pone davanti alla società come un compito pratico, non come un enigma logico.

Bollettino

Con questa rubrica, la redazione si propone di dare notizia di convegni, seminari e progetti di ricerca, ritenuti di particolare rilievo per le tematiche trattate dalla rivista.